IDIAP RESEARCH REPORT

Asynchronous detection and classification of oscillatory brain activity

Ricardo Chavarriaga [a] Ferran Galán [a b]
José del R. Millán [a c]

IDIAP–RR 08-36

May 2008

revised in February 2008

published in
16 European Signal Processing Conference (EUSIPCO-2008), August 2008

[a] IDIAP Research Institute, Martigny, Switzerland
[b] Universidad de Barcelona, Barcelona, Spain
[c] Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

IDIAP Research Institute **www.idiap.ch**
Av. des Prés–Beudin 20 P.O. Box 592 1920 Martigny – Switzerland
Tel: +41 27 721 77 11 Fax: +41 27 721 77 12 Email: info@idiap.ch

IDIAP Research Report 08-36

Asynchronous detection and classification of oscillatory brain activity

Ricardo Chavarriaga Ferran Galán José del R. Millán

May 2008

revised in February 2008

published in
16 European Signal Processing Conference (EUSIPCO-2008), August 2008

Abstract. The characterization and recognition of electrical signatures of brain activity constitutes a real challenge. Applications such as Brain-Computer Interfaces (BCI) are based on the accurate identification of mental processes in order to control external devices. Traditionally, classification of brain activity patterns relies on the assumption that the neurological phenomena that characterize mental states is continuously present in the signal. However, recent evidence shows that some mental processes are better characterized by episodic activity that is not necessarily synchronized with external stimuli. In this paper, we present a method for classification of mental states based on the detection of this episodic activity. Instead of performing classification on all available data, the proposed method identifies informative samples based on the class sample distribution in a projected canonical feature space. Classification results are compared to traditional methods using both artificial data and real EEG recordings.

1 Introduction and motivation

The characterization and recognition of electrical signatures of brain activity constitutes a real challenge. Applications such as Brain-Computer Interfaces (BCI) are based on the accurate identification of mental processes in order to control external devices (e.g. neuroprosthesis, wheelchairs or computers) without using the nervous system's efferent pathways [1]. Typically, users learn how to voluntary modulate different oscillatory EEG rhythms by the execution of different mental tasks [2]. In particular, asynchronous interfaces are required to identify the moment when the user provides meaningful information about his/her intentions or the underlying cognitive processes taking place in the brain.

Traditional BCI systems assume that neurological phenomena that characterize mental states is continuously present during their execution. However, it is known that endogenous mental processes like visual attention or object recognition are characterized by induced modulation of oscillatory activity that is not synchronized to any external stimuli (e.g. their latency may vary across trials) [3, 4]. As a consequence of this, recognition of these tasks may be improved by identifying the appearance of these episodic oscillations, and performing the classification of mental tasks based mainly –if not exclusively– on the activity during these periods.

Moreover, it has been proposed that synchronization of neural activity allows for efficient communication across brain areas required for cognitive processing in humans [3, 5]. Along these lines, Freeman used the term *frames* to describe active intermittent induced spatial patterns of amplitude modulation of beta-gamma oscillations in response to conditioned stimuli [6]. We borrow the term *frame* to denote episodic, discriminant oscillatory activity.

In this paper we describe a method for classification of mental states based on the detection of such episodic activity (i.e. frames) [7]. Opposed to other approaches previously proposed, no particular assumptions are made about latency of such episodes [8]. The proposed technique is validated and compared to traditional classification approaches using both artificial data and real EEG recordings.

2 Method description

The proposed method is composed of three stages. For every new sample we first extract features from the original signal (frequency domain) by using Canonical Variate Analysis (c.f. section 2.1). Then we determine whether that sample conveys discriminant information about the current mental state (i.e. it is a frame), based on the estimated sample distribution on the feature space (c.f. section 2.2). Finally, classification is performed only on those samples identified as frames.

2.1 Feature extraction: Canonical variate analysis

We compute relevant features using Canonical Variates Analysis (CVA), also known as Multiple Discriminant Analysis [9, 10]. This method provides canonical discriminant spatial patterns (CDSP) which direction maximizes the differences in mean spectral power between a given number of classes.

For a given class i, let $\mathbf{S_i} = \mathbf{s'_{i1}}, ..., \mathbf{s'_{in_i}}$ be a n_i x c matrix with the spectral estimated power of a given frequency band, where n_i is the number of samples and c the number of channels. Being k the number of classes, and $\mathbf{S} = (\mathbf{S'_1}, ..., \mathbf{S'_k})$, the $k-1$ CDSP of $\mathbf{S}$ are the eigenvectors $\mathbf{A}$ of $\mathbf{W^{-1}B}$ whose eigenvalues $\lambda_u, (u = 1, .., k-1)$ are larger than 0. Note that the direction of eigenvectors $\mathbf{A}$ maximize the quotient of the between-classes dispersion matrix.

$$\mathbf{B} = \sum_{i=1}^{k} n_i(\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})' \tag{1}$$

and the pooled within-classes dispersion matrix,

$$\mathbf{W} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (\mathbf{s}_{ij} - \mathbf{m}_i)(\mathbf{s}_{ij} - \mathbf{m}_i)' \tag{2}$$

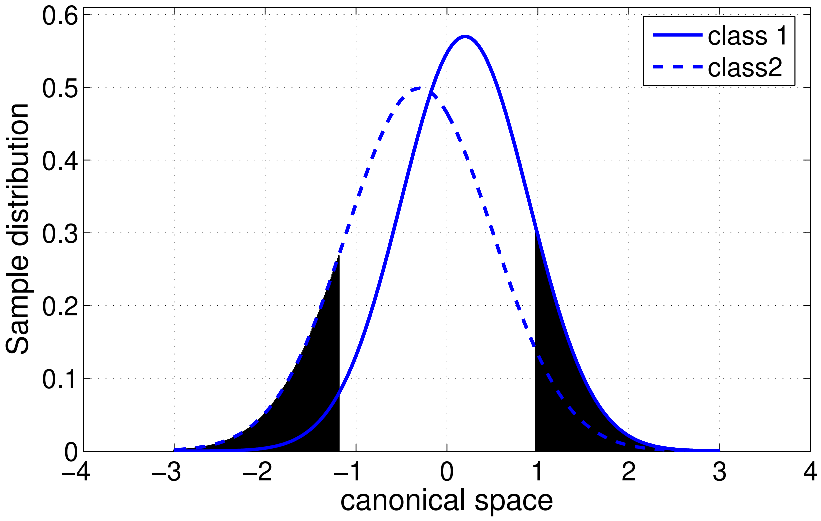

Figure 1: Class probability distribution in the canonical space. Only samples that do not lie in the overlapping region of both distributions ($y \notin [\theta_l, \theta_h]$) are considered as frames. In this example, the thresholds are set to percentile P_{10}.

where,

$$\mathbf{m}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{s}_{ij} \qquad \mathbf{m} = \frac{1}{n}\sum_{i=1}^{k} n_i \mathbf{m}_i \tag{3}$$

are the class and total centroids respectively. The CVA transformation yields a projection of the signal onto a space of $k-1$ dimensions, according to,

$$\mathbf{Y} = \mathbf{SA} \tag{4}$$

Interestingly, this method can also be used to rank the available channels given their contribution to the projected features on the new space. A measure of the discriminant power (DP) for each channel can be obtained based on the correlation matrix between the original channels and the new features in $\mathbf{Y}$ [10].

2.2 Frame detection and classification

As mentioned earlier, the proposed method relies on the detection of the most discriminative samples. We can estimate how discriminant a particular sample is based on the sample class distribution in the projected canonical space $\mathbf{Y}$. That is, samples lying in overlapping regions of the class distributions are less informative than those in non-overlapping regions.

Under this approach, we will first attempt to recognize informative phenomena, i.e. by identifying samples that lie on non-overlapping regions of the canonical space, and then perform classification based solely on those samples. If we consider a two-class problem, the canonical projection lies in a one-dimensional space and the overlapping region can be simply defined by two thresholds θ_l and θ_h. A projected sample y will be used for classification if and only if $y < \theta_l$ or $y > \theta_h$. A simple way to define these threshold is to use percentiles of the sample distributions. For instance, we will consider a sample as informative if it falls below a given percentile for class 1 (e.g. P_5) or above the opposite percentile for class 2 (e.g. P_{95}). In the rest of the document when referring to percentile P_n, we refer to both the percentile P_n, for one class, and percentile P_{100-n}, for the other class. The rationale of the proposed method is illustrated in Figure 1. Here the sample probability distributions for both classes are represented by Gaussians with different means.

Once a frame has been identified, classification is performed using Linear Discriminant analysis (LDA). According to our method, only those samples recognized as informative phenomena are used for both training and testing the performance, while other samples are ignored. In the next section, we compare the classification performance of the proposed method with the performance of classifiers trained and tested using all the projected samples (i.e. traditional BCI approach).

3 Experimental Results

3.1 Artificial data

3.1.1 Signal description

We validate the method using artificial data containing episodic oscillatory phenomena. We consider a simulated signal S that may contain two types of episodic phenomena (type I and type II). Both types contain similar frequency components –denoted as oscillations a, b and c– which may occur with different probabilities. The oscillatory components are described in the Table 1.

Oscillations	**Frequency**	**Duration**
a	55Hz	200ms
b	80Hz	100ms
c	30Hz	100ms

Table 1: Oscillatory components used in the artificial data

Episodes of the different phenomena are composed of the sum of the three oscillations, whose amplitude U is probabilistically determined as follows,

- Events Type I:
$P(U = 1 \mid b) = 1$
$P(U = 2 \mid a) = 0.4$ and $P(U = 0 \mid b) = 0.6$
$P(U = 2 \mid c) = 0.4$ and $P(U = 0 \mid c) = 0.6$

- Events Type II:
$P(U = 1 \mid a) = 1$
$P(U = 2 \mid b) = 0.4$ and $P(U = 0 \mid b) = 0.6$
$P(U = 2 \mid c) = 0.4$ and $P(U = 0 \mid c) = 0.6$

Ten different trials were generated with a length of 90s. Every trial is composed by asynchronous episodes of both types. Each episode lasts 1 second, and the interval between two successive episodes varies randomly between 1 and 3 seconds (uniform distribution). Both types of episodes have the same probability of appearance. Besides the episodic phenomena, white noise is added to the entire trial length. Figure 2, shows an example the signal S.

Finally, we assume three different sensors E1, E2, and E3. Sensor E1 captures the signal S with no distortion (E1 = S). Signal captured by E3 is only composed by white noise (E3 = N), while E2 captures a combination of the previous two (E2 = $\eta E1 + (1-\eta)E3$, $\eta = 0.2$). Given that the CVA method ranks the channels according to their discriminant power [10], we expect the extracted features to be highly correlated with channel E1, while channel E3 should contribute the least to the resulting canonical projection.

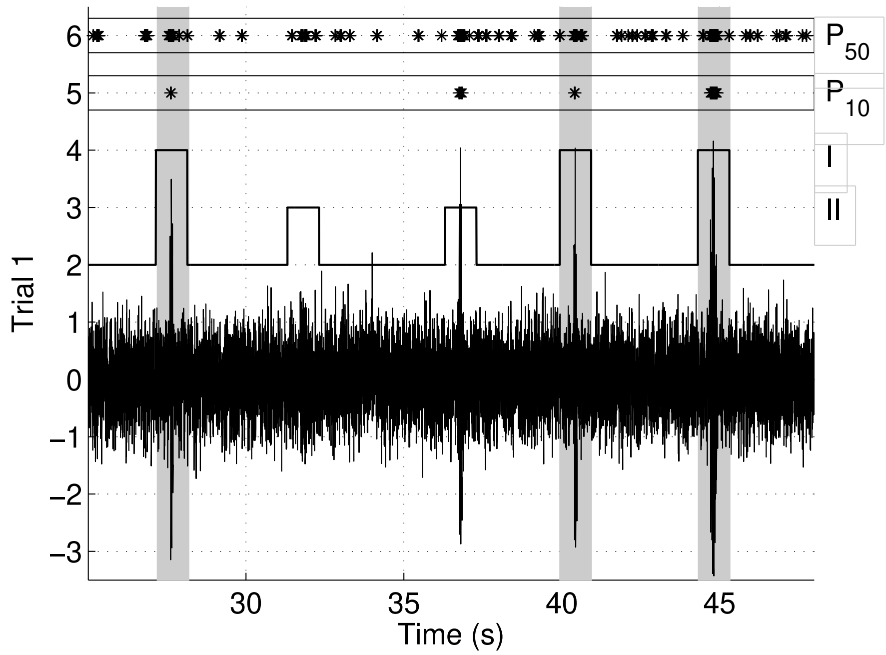

Figure 2: Example of the artificial signal S (fragment of a single trial). Episodes of the different phenomena (I and II) are shown in the middle part of the plot. Time samples identified as frames of Type I episodes using percentiles P_{10} and P_{50} are marked with asterisks (top).

3.1.2 Frame detection and classification

We test the ability of the method to recognize and correctly classify frames associated to episodes of type I. It should be noticed that although the length of each episode is 1 second, the characteristics oscillations for both types of events last 200 ms at most. Therefore, some samples labeled as type I bear no difference with samples where no event is present.

For the three sensors (E1, E2, E3) we extract the continuous Morlet wavelet coefficients on 13 frequency components (12, 28, 32, 36, 40, 44, 48, 56, 64, 72, 80, 88, and 96Hz). Separate canonical projections and classifiers are built for each frequency component. The threshold for frame recognition is set to percentile P_{10}. We assess single-trial classification and generalization by means of 10-fold cross-validation where each fold corresponds to a separate trial.

The discriminant power of each channel is shown in Figure 3. As expected, channel E1 contributes the most to the canonical projection in all frequency components, while the signal on channel E3 is not relevant.

We identify the most informative samples according to the sample probability distribution in the projected canonical space (i.e. frames). Figure 2 shows an example of the identified frames for thresholds set at percentiles P_{10}, and P_{50}. It can be noticed that a less restrictive threshold (i.e. P_{50}), increases the number of detected frames that do not correspond to relevant events. In contrast, more restrictive thresholds yield better frame recognition, at the expense of a smaller number of detected frames.

The next step is to recognize how many of the detected frames are correctly classified as belonging to episodes of type I. Classification accuracies are shown in Figure 4 for both the traditional and proposed methods. Random classification accuracy is obtained when using all trial samples to train and test the classifier (i.e. traditional BCI approach). Such a poor performance may be explained by inconsistencies on the labeling of train trials as explained above, as well as the probabilistic nature of oscillations that characterize both relevant and irrelevant events. When classifying only the frame samples, classification accuracy increases significantly, with a maximum average recognition rate above 85% (86.19% and 85.71% for 72Hz and 80Hz respectively), corresponding to the oscillation type b, consistently present on the interesting phenomena (type I events).

3.2 Real EEG data

3.2.1 Experimental protocol and preprocessing

We record EEG data from two subjects (one female) using a portable Biosemi acquisition system using 64 channels (10/20 international system) sampled at 512 Hz, and high-pass filtered at 1Hz. Subjects seat comfortably in front of a computer screen and fixate their gaze at a cross placed at the center of the screen. Upon request of the experimenter, they covertly attend one of two possible spatial targets (i.e. lower-left and lower-right corners of the monitor). Targets are selected in a pseudo-random balanced order. The experiment consists of 10 recording sessions, composed by 4 trials each (two trials per target). Every trial lasts 7 second, but only the first 600ms of each trial are kept for the analysis.

At each time point, EEG signal is re-referenced to the mean activity over all electrodes (i.e. common average reference, CAR). We compute 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 96Hz) for 16 electrodes shown in Figure 5. Selected electrodes were chosen based on a preliminary analysis of the wavelet coefficient scalp topography. According to this, each trial is composed by 512x0.6 samples and 18x16 features.

The motivation of this paradigm is to explore the potential use of voluntary modulation of visual attention in Brain-Computer Interfaces. For example, in navigation tasks, to shift visual attention –without eye movements– towards the intended direction of movement is more natural to the user than the mental tasks frequently used in this applications (e.g. motor imagery, arithmetic tasks). Previous studies have used visual attention in steady-state visual evoked potential (SSVEP) paradigms [11]. However this approach relies on external cues, and is not suitable for asynchronous operation of BCI systems. Recent studies have shown modulations of the EEG alpha-band due to changes in visuo-spatial attention [12], as well as changes in the gamma band corresponding to endogenous shifts of attention [4]. These studies support the idea of recognizing oscillatory activity as a marker for changes in visual attention.

3.2.2 Frame detection and classification

The goal is to classify the spatial location the subject is attending to (i.e. left or right). As described in section 2, a canonical projection is built for each frequency band, and the sample probability distribution allows to detect the frames to be used for recognition. In the following results we set the threshold for frame identification to the percentile P_5.

Classification is performed using Linear Discriminant Analysis (LDA) using both the traditional and the frame approach. For each subject we assessed the performance using k-fold cross validation ($k = 20$), where each fold was composed of one trial of each condition, respecting the original timing when they were recorded.

Figure 6 shows the classification performance for each frequency band using the traditional approach, i.e. using all projected samples for both training and testing. Maximum average performance is 58.41% at 10Hz, and 63.08% at 12Hz for subject 1 and 2 respectively. In contrast, using the frames recognition method, shown in Figure 7, the average classification accuracy is 80.46% at 72Hz, and 87.31% at 32Hz for subjects 1 and 2 respectively. This is consistent with previous studies that have reported modulation of Gamma-band oscillations (>30Hz) as a result of endogenous shifts of attention [4]. Note that this performance is computed only on those samples recognized as frames.

This method relies on the idea that underlying neurological phenomena is not necessarily present in a continuous manner, and some samples may be more informative than others. Accordingly, trial-based classification can be based on a subset of samples. Figure 8 shows examples of test trials projected onto the canonical space. In this plot, time samples recognized as frames are denoted by asterisks, while other samples are neglected as not being discriminative enough. The fact of rejecting these samples allows a ten-fold increase in the theoretical information transfer rate (i.e. channel capacity [1]) with respect to the traditional approach [7].

4 Discussion

We have presented a method for mental states classification based on the asynchronous detection of discriminant samples. Contrasting to traditional approaches where classification is performed on all the available data, we hypothesize that discriminant phenomena is not uniformly present throughout the whole execution time. Experimental results on both synthetic and real data seem to confirm that not all samples are equally informative about the mental state of the subject. In both cases the detection of informative samples (i.e. frames) yields better classification performances. It should be noticed that the proposed methodology is general-purpose and can be used in applications different than classification of EEG signals (c.f. tests on artificial episodic data).

Under the assumption that only a subset of those samples convey discriminant information, the fact of using all samples for building BCI classifiers implies that those classifiers are trained using both relevant and irrelevant information, which may affect the overall system performance. Conversely, when trying to decode mental tasks characterized by episodic activity, it would be wrong to expect a sustained classification output throughout the whole duration of each trial. Some BCI implementations have dealt with this fact by rejecting samples (i.e. classifying them as *unknown*) whenever there is not enough confidence on the classifier output [1]. The proposed approach, tries to solve this problem by classifying only those samples identified as being discriminative (c.f. Figures 2 and 8).

In this study we used a simple thresholding criteria to define which samples are taken into account for classification. In the future we will explore more formal ways to define those thresholds based on the sample probability distributions estimated from training data. Assuming that we have good estimations of these distributions, thresholds can be defined following probabilistic criteria in order to maximize the confidence of a sample corresponding to a particular class.

Finally, it should be noticed that the classification approach used in this work classifies each time sample in an independent way. Previous studies have proposed the combination of independently classified samples –through accumulating evidence through time, or averaging over consecutive classified samples –as a way to improve the reliability of the classification [1, 13]. Such methods might increase the robustness of the current approach for event recognition.

5 Acknowledgments

This work has been supported by the Swiss National Science Foundation NCCR-IM2 and by the EC-contract number BACS FP6-IST-027140. This paper only reflects the authors' views and funding agencies are not liable for any use that may be made of the information contained herein.

References

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

[2] N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen, B. Kotchoubey, A. Kübler, J. Perelmouter, E. Taub, and H. Flor. A spelling device for the paralysed. *Nature*, 398:297–8, 1999.

[3] C. Tallon-Baudry and O. Bertrand. Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn Sci*, 3:151–162, 1999.

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

[5] S. Ray, E. Niebur, S.S. Hsiao, A. Sinai, and N.E. Crone. High-frequency gamma activity (80-150hz) is increased in human cortex during selective attention. *Clin Neurophysiol*, 119:116–133, 2008.

[6] W.J. Freeman. Origin, structure, and role of background EEG activity. Part 2. Analytic phase. *Clin Neurophysiol*, 115:2089–2107, 2004.

[7] F. Galán, J. Palix, R. Chavarriaga, P.W. Ferrez, E. Lew, C.A. Hauert, and J. del R. Millán. Visuo-spatial attention frame recognition for brain-computer interfaces. In *Int Conf Cognitive Neurodynamics*, 2007.

[8] F.B. Vialatte, C. Martin, R. Dubois, J. Haddad, B. Quenet, R. Gervais, and G. Dreyfus. A machine learning approach to the analysis of time-frequency maps, and its application to neural dynamics. *Neural Netw*, 20:194–209, 2006.

[9] W.J. Krzanowski. *Principles of multivariate analysis*. Oxford University Press, Oxford, 1998.

[10] F. Galán, P.W. Ferrez, F. Oliva, J. Guàrdia, and J. del R. Millán. Feature extraction for multi-class BCI using canonical variates analysis. In *IEEE Int Symp Intelligent Signal Processing*, 2007.

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

[12] G. Thut, A. Nietzel, S.A. Brandt, and A. Pascual-Leone. Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci*, 26:9494–502, 2006.

[13] S. Lemm, C. Schäfer, and G. Curio. Aggregating classification accuracy through time - classifying single trial EEG. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Adv Neural Information Processing Systems 19*, pages 825–832. MIT Press, Cambridge, MA, 2007.

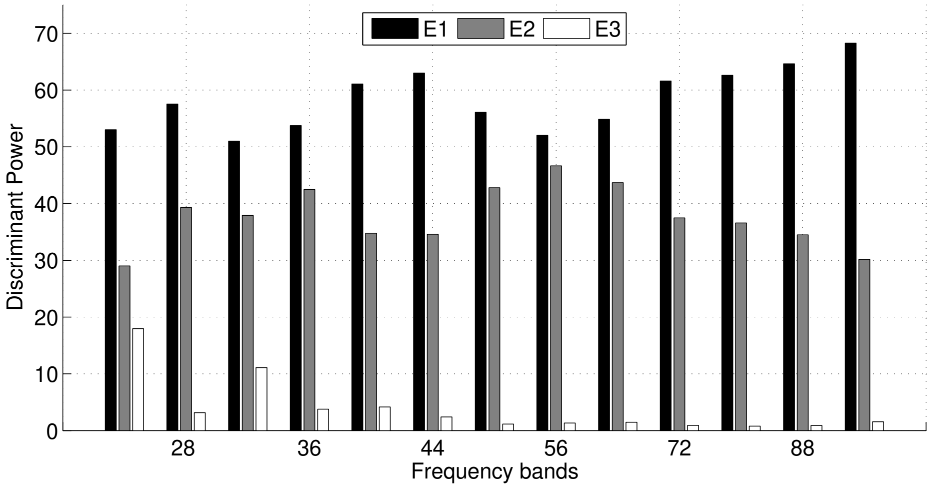

Figure 3: Average discriminant power (DP) of each channel for the different frequency bands.

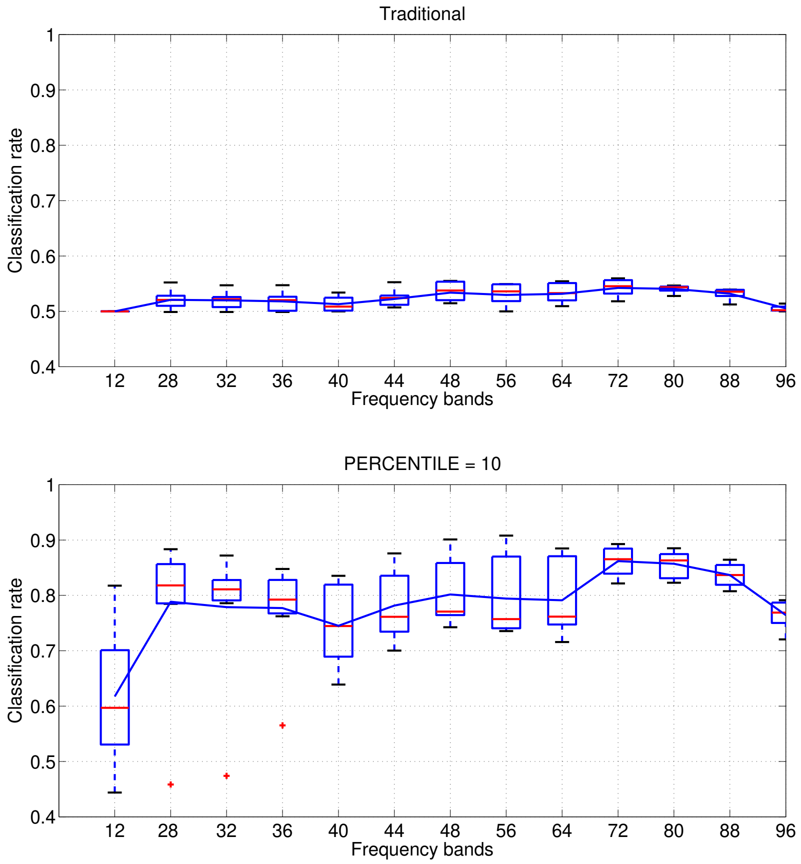

Figure 4: Classification performance on artificial signals. *Top*, Traditional approach, classifiers are trained and tested using all samples of each trial. *Bottom*, Classification is only performed on samples identified as frames. The threshold is set to percentile P_{10}.

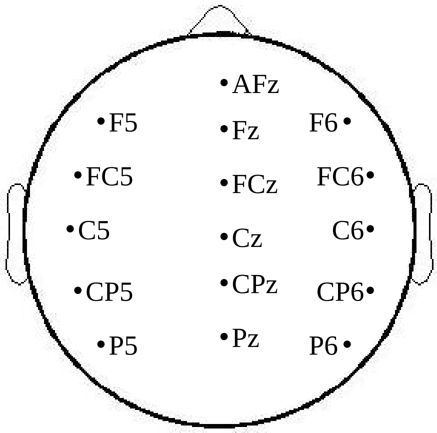

Figure 5: Electrodes used for the recognition of visuo-spatial attention modulation: F5, FC5, C5, CP5, P5, AFz, Fz, FCz, Cz, CPz, Pz, F6, FC6, C6, CP6, P6.

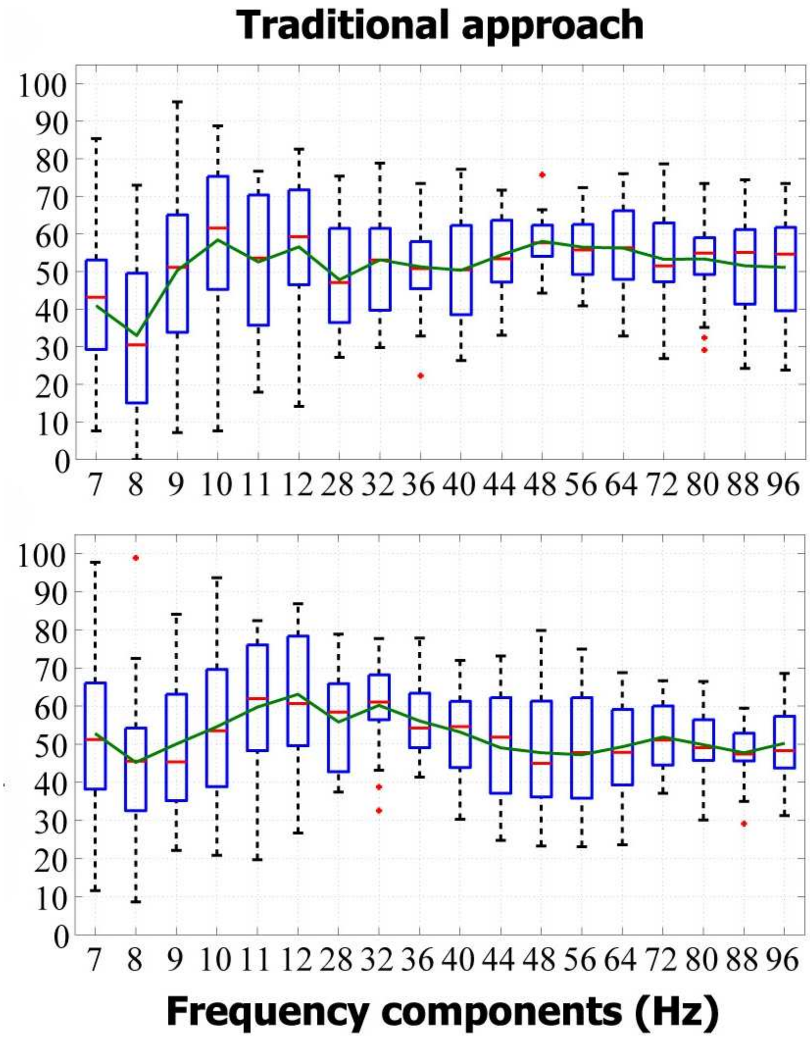

Figure 6: LDA Classification accuracy (k-fold cross validation) on real EEG data using the traditional approach. X-axis corresponds to the 18 frequency components used in the study. *Top:* Subject 1. *Bottom:* Subject 2.

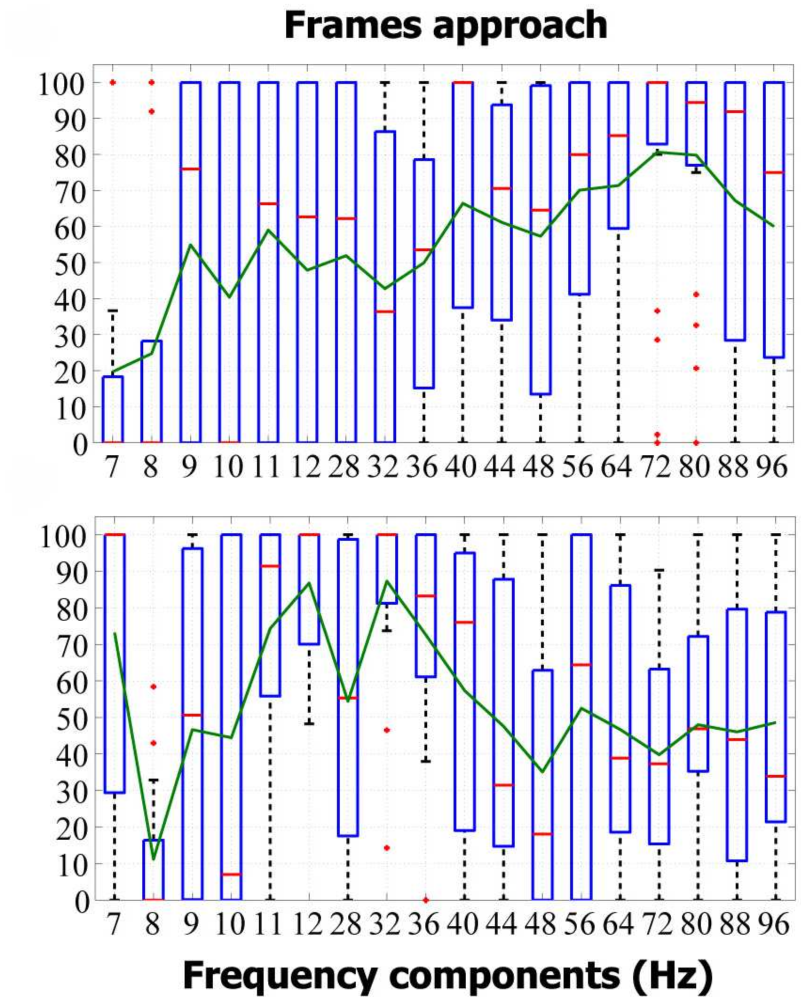

Figure 7: LDA Classification accuracy (k-fold cross validation) using the proposed approach. X-axis corresponds to the 18 frequency components used in the study. *Top:* Subject 1. *Bottom:* Subject 2.

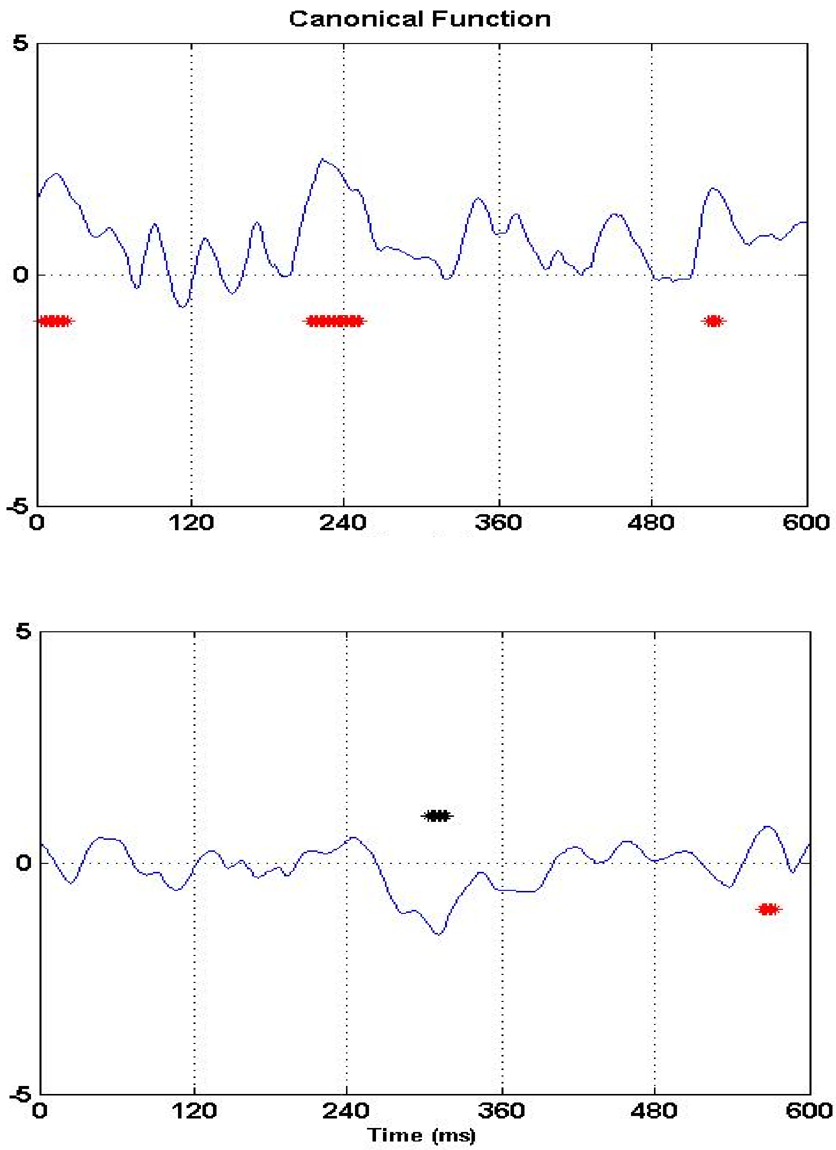

Figure 8: Detection of informative samples. Each plot shows the canonical projection of a test trial. Identified frames are marked as asterisks. The upper plot shows a trial where all the frames are correctly classified (i.e. all samples identified as frames correspond to the same, correct class). The bottom plot shows a trial where some frames are misclassified.