U.P.B. Sci. Bull., Series C, Vol. 80, Iss. 1, 2018 ISSN 2286-3540

ENHANCED CLASSIFICATION METHODS FOR THE DEPTH OF COGNITIVE PROCESSING DEPICTED IN NEURAL SIGNALS

Irina-Emilia NICOLAE[1,2], Laura ACQUALAGNA[2], Georgeta-Mihaela Neagu (UNGUREANU)[1], Rodica STRUNGARU[1], Benjamin BLANKERTZ[2]

Analyzing brain states is a difficult problem due to high variability between subjects and trials, therefore improved techniques are requested to be developed for a better discrimination between the neural components. This paper investigates multiple enhanced classification methods for neurological feature selection and discrimination of the depth of cognitive processing. The aim is to detect the strengths and weaknesses of different classification methods and benefit from their highest performances, so that the neural information could optimally be detected. As a result, we obtained a classification rate improved by at least 5% by integrating complementary information that better describe the neural activity.

Keywords: Electroencephalography (EEG), signal processing, classification, cognitive processing

1. Introduction

Brain signal processing is a core research nowadays and powerful methods are developed for machine learning and signal processing. We focus on evaluating the main brain investigation techniques, considering especially the popular methods used in EEG signal processing [1], considering the temporal [2] and spectral classification [3, 4] for detecting the ERPs and the oscillatory activity.

A crucial step in brain signal processing is to enhance the differences between the neural correlates of the desired actions. It is mandatory for a feature extraction mechanism to detect the corresponding components and avoid the unwanted artifacts, especially when it comes to higher order data. The fundamental solution consists in selecting the components and features that describe the neuronal sources in the best way. Differently from the majority of Brain-Computer Interfaces research, where only the effect of visual attention is investigated, we explore the users' modality of perceiving the external information by inspecting different degrees of processing. This will push us closer

[1] Dept. of Applied Electronics and Information Engineering, University POLITEHNICA of Bucharest, Romania, e-mail: irina.nicolae@aut.pub.ro; g.mihaela.ungureanu@upb.ro; rodica.strungaru@upb.ro;

[2] Neurotechnology Group, Technische Universität Berlin, Berlin, Germany

to estimate user's internal processes of intention and decision regarding the outside information.

Different classification methods and scenarios have been investigated and applied to extract the neural correlates of the depth of cognitive processing and their performances were evaluated based on the accuracy values. In this manner, the effectiveness of the proposed methods is tested, and we chose the one that quantifies better the way users process the presented information. On this line, we explore different methods of temporal and spectral classification to discriminate different levels of cognition regarding the visual information processing, namely the temporal ERP methods [5] and the Common Spatial Patterns technique for the spectral domain [6, 7]. Regarding the classification approaches, we referred to the feature detection step to improve the classification performance. We evaluated the classification considering each domain separately, with fixed temporal windows and sliding window for the temporal domain and within different frequency bands for the spectral domain, and also in a combined tempo-spectral approach. We then analyzed and compared their performances highlighting their strength points and disadvantages.

2. Classification methods

1) Temporal classification

The effects of visual processing are expressed by the Event Related Potentials (ERP) which are observed in the temporal evolution of the brain signals [2]. For that, it is necessary to investigate and extract the temporal features. Among the general used temporal methods [2], we chose the ERP detection techniques which have the best classification performance, as shown in [2]. In addition, to a fixed window ERP classification [8], we have investigated a sliding window approach [9], in order to overcome the temporal variability. The applied temporal methods will be described below, and a comparison is provided in the results section.

a) ERP classification

The algorithm proposed by [6] and applied in Nicolae et al. [8], detects suitable temporal windows by heuristic selection of five-time intervals showing the highest and constant discrimination between classes, based on the point biserial correlation coefficient, *signed* r^2.

$$\text{singed r}^2 = \operatorname{sgn}\left(\left(\frac{\mu_1 - \mu_2}{\sigma}\sqrt{\frac{n_1 n_2}{n^2}}\right)^2\right), \tag{1}$$

where μ_1 – mean of class 1, μ_2 – mean of class 2, σ – standard deviation, n_1 – number of data points of class 1, n_2 – number of data points of class 2, and n – the entire sample size.

In addition, the brain activity is generated in different brain regions, and so the spatial distribution of the electrical potentials has also to be considered. We consider the features in ensemble, by conjointly taken both the spatial and the temporal features [5]. The search was performed on the trial period together with the relaxation period that followed, having 2 s in total. Additionally, when performing the spatial improvement, all the available channels are taken into account while the classifier establishes an optimal weighting of the channels to discriminate the brain activity.

b) Sliding ERP classification

The classifier is trained in advance on the most discriminant time intervals (Table 1) selected for each condition and discrimination class, given by the signed r^2 differentiation shown in [8], and afterwards applied in sliding manner, with 50 ms windows length and 10 ms time shifts starting from the first time point of each trial and exploiting all the available channels. The evaluation is performed trial by trial, with cross-validation, as described in the '*Classification evaluation*' sub-section.

Table 1

Discriminant time intervals selected for each condition and class pair

Class pair	Time interval (ms)		
	Memory	Language	Visual imagery
NP – SP	400 - 800	300 - 900	350 - 500
NP – DP	300 - 800	350 - 850	350 - 850
SP - DP	300 - 550	750 - 1050	350 - 650

2) Spectral classification

The effects of cognition are observed also in the spectral activity; therefore, we investigate the neural activity considering the spectrum up to 50 Hz.

a) Spatio-Spectral Decomposition, SSD

Spatio-Spectral Decomposition, SSD [10] was tested and performed to increase signal to noise ratio and so to enhance the classification performance.

SSD detects individual activity sources relating to high variance in the frequency of interest and low variance for the neighboring frequency and reduces the redundant components caused by volume conduction. This linear spatial filtering prior to Common Spatial Pattern method is very useful when optimizing the discrimination of the power spectrum peaks. The considered selected time interval, containing the cognitive processing, was up to 350 ms after the stimuli.

b) Single band spectral classification

For reducing the effect of volume conduction, Common Spatial Pattern (CSP) [6, 7] was applied. CSP enhances the detection of the brain states by spatial filtering which increases the signal of interest and suppresses the background activity. CSP's objective is to maximize the variance for a class while minimizing the variance for another class, in a binary selection. Three optimal components were selected per class, given 6 spatial filters for each binary discrimination.

Based on the highest discrimination detected in the spectrum activity, we evaluated different frequency bands for the classification: θ (5-7 Hz), α (8-14 Hz) and β (16-20 Hz) frequency bands.

c) Dual band spectral classification

We combined the spectral features considering the most relevant frequency bands that show highest discrimination. Therefore, we integrate the α (8-14 Hz) and β (16-20 Hz) band power features to consider a wider range for the neural activity components.

*3**) Combined temporal and spectral classification***

For a better estimation of user's cognitive processing level, we consider also integrating the temporal features selected with the signed r^2 discrimination and the spectral features given by the dual band CSP method, further described in Ref. [11]. In order to perform the combined single trial classification, the temporal ERP features were concatenated with the CSP dual-band features. The temporal features of 5×62 features, consist of 5 features for the 5-time intervals selected on each of the 62 channel locations. The oscillatory features are represented by 6 features for each band as logarithm of the variance of the spatial filters selected by CSP, containing three features for each class of the binary pair. Forward, this high number of features was used for the shrinkage LDA classifier.

4) Classification evaluation

Considering the temporal methods previously described, supervised classification was performed referring to the stimuli labels by using regularized Linear Discriminant Analysis, LDA, with shrinkage [2, 5, 7]. The evaluation was carried out in 10 folds cross-validation with 10 repetitions, meaning that the data was split in 10 parts, nine for training and one for testing and the procedure is repeated 10 times by changing the sets. The results are then averaged to obtain a single estimation. The performance of the classification is assessed by the area under the ROC (Receiver Operator Characteristics) curve, where the ROC curve is obtained by plotting the False Positive Rates (on the abscissa) against the True Positive Rates (on the ordinate) and the area computed under this curve (the AUC value) represent the measure of classification performance, reaching a maximum one. The grand average error is estimated by the standard error of the mean across participants, *sem*.

$$sem(x) = \frac{\sigma(x)}{\sqrt{n}}, \tag{2}$$

where

sem(x) – standard error of the mean for data *x*;
σ(x) – standard distribution of the data *x;*
n – number of participants.

For the validation of the spectral classification scenarios, as proposed in [6], the relevant CSP components were detected from the training data and used to spatially filter the test data, accordingly. As classifier, LDA was applied and the procedure was similarly performed in cross-validation form, with 10 folds and ten repetitions. The classification assessment is likewise analyzed by the area under the ROC curve.

3. Experimental study

The varied classification approaches were tested on a BCI study, with 15 participants, in which the subjects viewed a sequence of visual stimuli with cartoons drawing representations of fruits, animals and transportation devices, and performed different tasks requested by each type of stimuli. Briefly, three levels of cognitive processing were investigated, namely: no-processing, shallow processing, which required stimuli differentiation based on color, and deep processing, which imposed the performing of a cognitive process. In the last case, the evaluated cognitive processes were memory, language and visual imagery. The memory process needed visual comparisons to a previous memorized stimulus. The language process demanded semantic correlations based on the words that represent the presented images. And finally, the visual imagery condition involved mental imagery of the respective objects in reality and performing dimension comparisons. In total, 600 trials of 1.25 s long were presented divided in 75% for the no-processing class (NP) and 12.5% for each shallow (SP) or deep processing class (DP). For the signal processing we considered also the relaxation period that follows the trail period of 0.75 s length, due to the extended execution of the complex cognition (more details about the study are provided in [11, 12]). The data were collected at Technische Universität Berlin and the experiments were approved by the Ethical Review Board of the University.

The behavioral data were evaluated considering the absolute difference between participant response and the correct number, divided by the correct number. This gives a value between 0 and 1, where a value close to zero represents higher performance. The final assessment value obtained for each condition and participant considers the averaged performances for all runs, as in eq. 3.

$$a_{p,C} = \frac{\sum_{r=1}^{5} \frac{pr_r - c_r}{c_r}}{5}, \tag{3}$$

where

$a_{p,\,C}$ – the assessment value for the behavioral data of the participant p ($p \in \overline{1-15}$), for the condition C (memory, language or visual imagery)

pr_r – the participant's response for the run r, $r = \overline{1-5}$

c_r – the correct number for the run r

4. Results

We analyzed the single trial classification performance of each separate classification scenario, by comparing them and highlighting the approaches that led to the ensemble classification approach, ERP and multi band CSP with SSD, described in [11]. The classifiers were binary applied among the three classes: no-processing, shallow processing and deep processing.

As it can be seen in the figure 1, the CSP classification on the 5-7 Hz frequency band (in dark yellow) gives the lowest performance among all the classification approaches. An improvement is observed for the 8-14 Hz (yellow) and 16-20 Hz (orange) frequency bands, which were selected also for the combined classification. For these most discriminative frequency bands, α and β, we evaluated the effect of the Spatio-Spectral Decomposition algorithm (light green and dark green) which shows, in general, a better improvement for the performance of the classification. More exactly, it enhances the classification more emphasized for the SP-DP discrimination of the 8-14 Hz frequency band (light green). Next, the multi-band classification based on SSD which considers both 8-14 Hz and 16-20 Hz frequency bands (red), shows a small improvement when discriminating the deep processing class. Overall, all methods provide classification greater than the chance level of 50%. Considering the temporal classification based on ERP with fixed windows (dark blue) in comparison with the spectral approach based on multi band CSP and SSD (red), it offers increased performance when discriminating the non-processing class, but decreased classification for the SP-DP pair. Therefore, by applying the combined method (magenta) which joints this ERP with fixed window based on signed r^2 with the multi-band CSP with SSD classification [11] and by taking into account both temporal and spectral features, an enhanced classification was obtained, even for the SP-DP pair. The ERP classification with sliding window (grey) shows lower performance, around 70% (more detail in Figure 2 and Figure 3), as compared to the ERP fixed window method with interval selection based on signed r^2 (dark blue) that reaches better results: classification rate is greater than 80% for NP-SP and NP-DP discrimination.

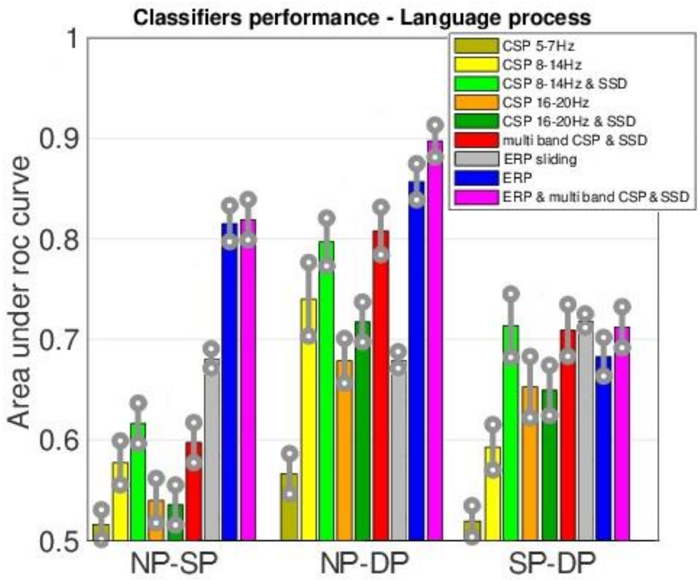

Fig. 1. The classifiers performance for the language condition considering the grand average AUC values (across all participants). The standard error of the mean (*sem*) is shown in grey.

For more details about the sliding ERP classification performance, we can visualize the LDA scores and losses of the classifier for each temporal window, where the maximum classification output within each temporal window was selected. Here, the scores are represented by the absolute value of the distance of each point or sample to the separating hyperplane. Below in figure 2 and figure 3 we can view the scores and losses respectively for the SP-DP classification.

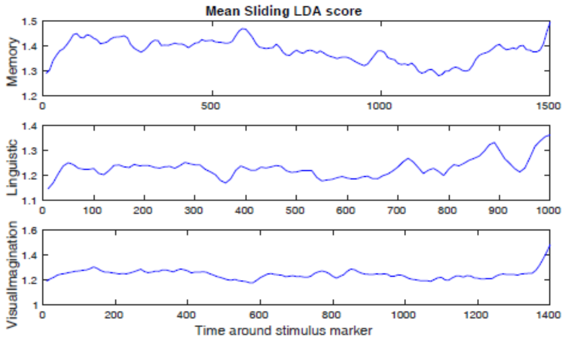

Fig. 2. The sliding ERP scores for the SP-DP classification

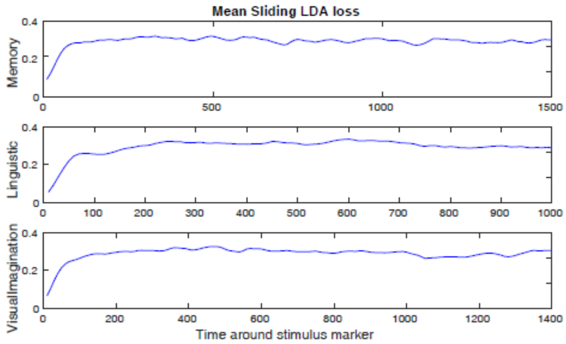

Fig. 3. The sliding ERP losses for the SP-DP classification

Finally, the average of the output classification of all windows and over all epochs was selected for the mean presented in Fig. 1, giving 0.69 AUC values ±0.016 *sem* for NP-SP classification, 0.68 AUC ±0.012 *sem* and 0.71 AUC ±0.0010 *sem* for the SP-DP classification, considering the language condition.

Now, for a more detailed visualization of the highest classification performance considering all classes, we analyze the distribution of the area under the ROC curve values (AUC) of the combined classification (ERP and multi band CSP with SSD [11], described in Section 3.3) among all conditions and within subjects (Table 2). As it can be observed in the table 1, all the values are beyond 56%.

Table 2

ERP and multi band CSP with SSD classification performance over all conditions and subjects

Subject	Memory			Language			Visual imagery		
	NP-SP	NP-DP	SP-DP	NP-SP	NP-DP	SP-DP	NP-SP	NP-DP	SP-DP
P1	0.6361	0.8198	0.7499	0.8282	0.8452	0.7507	0.6885	0.9484	0.7312
P2	0.8558	0.9359	0.7330	0.8171	0.7980	0.6621	0.8514	0.8735	0.6728
P3	0.8342	0.8819	0.7889	0.8630	0.8016	0.8139	0.7730	0.8935	0.6466
P4	0.8087	0.9501	0.7572	0.7943	0.9043	0.7827	0.8258	0.9097	0.7846
P5	0.7165	0.8220	0.6894	0.7703	0.9657	0.7787	0.8143	0.9304	0.7198

P6	0.7571	0.8792	0.7062	0.8212	0.9553	0.7508	0.6587	0.8321	0.6353
P7	0.6299	0.6160	0.6497	0.7907	0.8880	0.7197	0.7388	0.8729	0.7177
P9	0.7341	0.8040	0.6678	0.7350	0.8329	0.6539	0.7380	0.8541	0.7630
P10	0.7650	0.8556	0.7945	0.8502	0.8735	0.8040	0.8199	0.9194	0.7820
P12	0.6983	0.9165	0.7485	0.6574	0.9728	0.5616	0.7488	0.8878	0.6921
P13	0.8742	0.8714	0.7047	0.9073	0.9534	0.6455	0.8588	0.9306	0.6888
P14	0.7613	0.9524	0.7382	0.9432	0.9787	0.6920	0.9091	0.9581	0.7363
P15	0.9201	0.8095	0.6508	0.9254	0.8531	0.5755	0.8941	0.7287	0.6356
P16	0.7278	0.9133	0.7457	0.7572	0.9045	0.7151	0.5858	0.9386	0.7178
P17	0.7731	0.9362	0.6501	0.8600	0.9333	0.7767	0.7825	0.9235	0.7116
Mean	**0.7661**	**0.8643**	**0.7183**	**0.8214**	**0.8974**	**0.7122**	**0.7792**	**0.8934**	**0.7090**
sem	0.0213	0.0222	0.0126	0.0195	0.0159	0.0204	0.0230	0.0150	0.0124

Moreover, the classification performances among conditions and pair classes were statistically tested by ANOVA model ($p < 0.001$, $F = 64.99$). The analysis of variance was performed on the null hypothesis that the data comes from a standard normal distribution and was rejected at the 2% significance level ($p=0.02$). The classification performance shows a normal distribution observable in Figure 4, which represents that the classifier results do not present high variability between subjects and trials. For verifying that each group comes from a normal distribution with a mean and variance estimated from each group, we used the chi-square goodness-of-fit test and all tests passed at the 5% significance level. However, the groups were not statistically confirmed to follow a standard normal distribution using one-sample Kolmogorov-Smirnov test at the 5% significance level.

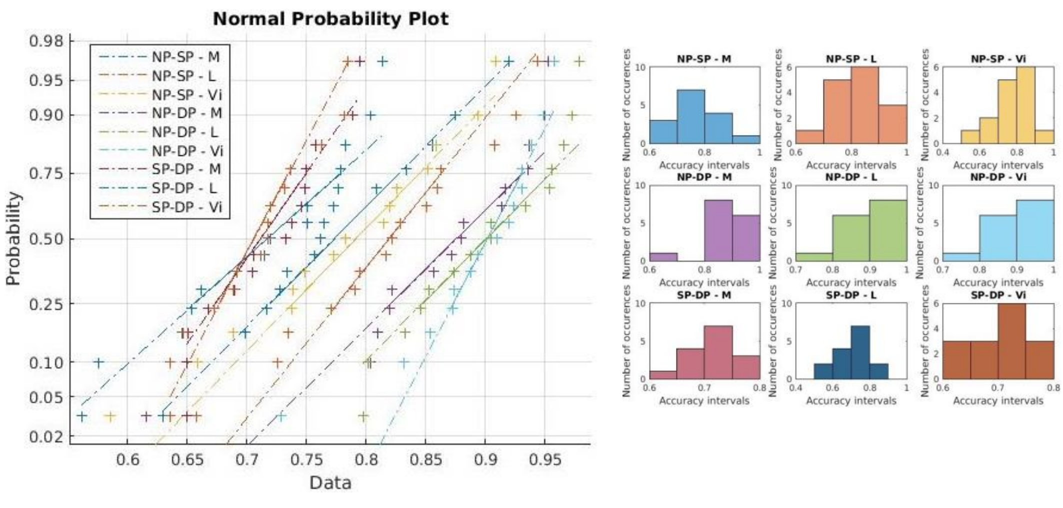

Fig. 4. The normal distribution graphs of the combined classifier performance (ERP and multi band CSP with SSD). The normal probability plot (left) and the histogram of each class (right).

We observed no significant difference for the classification performance among subject groups considering the experience in BCI. Although we noticed small differences considering behavioral results [8], this was not statistically significant (Table 3). The statistical tests were performed using two-sampled t-test at the 5% significance level. We divided the subjects in three categories based on their behavioral performance. First group shows accurate results with an assessment value between 0 and 0.05 given by 33.33% of the subjects. The second group was selected with assessment values between 0.05 and 0.1, representing 46.67% of the subjects, and the third group with values between 0.1 and 0.2, composed by 20% of the subjects.

Looking into the Table 3, over the results for the language condition, we observe a classification performance of at most 7% between the groups. Also, we observed a higher classification performance of at most 5% for the first group compared to the entire non-grouped data for the language condition in Table 2. This shows that an improvement regarding classification is consistent with the behavioral levels, but not statistically representative (at the 5% significance level using two-sampled t-test). We selected the language process for visualizing the classification results, because it showed the highest performance among all conditions. For the other conditions, the trend is similar with group 1 having the highest classification performance.

Table 3

ERP and multi band CSP with SSD classification performance between subjects' groups based on behavioral results for the language condition

Subject groups	Language		
	NP-SP	NP-DP	SP-DP
Group 1 (5, 6, 9, 11, 12)	0.8584	0.9453	0.7342
sem	0.0307	0.0185	0.0290
Group 2 (1, 3, 4, 7, 10, 13, 14, 15)	0.8095	0.8879	0.7120
sem	0.0286	0.0190	0.0334
Group 3 (2, 8)	0.7761	0.8155	0.6580
sem	0.0335	0.0142	0.0033

Furthermore, we detected no significant statistical difference considering gender, or participants' qualification, neither for the experiments performed in mother tongue or for the ones involving the second spoken language. Definitely, a larger study is necessary to confirm or infirm all these observations.

5. Conclusions

Our research paper compares multiple classification methods and identifies the most efficient ones that combine different features to improve the distinction between different neural characteristics and to provide higher accuracy levels. The temporal classification helps to discriminate the ERP characteristics between different levels of cognitive processing and the spectral classification, given by CSP helps to detect the neural patterns referring to the oscillatory activity. The extracted the spatial components representative to the alpha frequency band relate to attention and an easier processing, while complementary the beta spectral components represent more complex mental activity, as shown by Buzsáki in different ways [13]. By combining the temporal and the spectral classification, complementary information is taken into considered for a better estimation of the neural activity as shown by the classification outcome. As presented in the *Results* section, the temporal classification with sliding window did not provide significant accuracy on this data, which could mean that there are no significant fluctuations in time that could be detected on a 50ms window. Moreover, when considering the sliding LDA approach, the features represent the entire epoch time which on average, results in a lower accuracy (In the beginning and at the end of the trial there is no significant discrimination between classes). Therefore, the fixed window based on the highest discrimination interval between classes gives higher performance and this is the reason why we selected these temporal features for the combined approach.

As a further step in this manner we propose combining the ERP and spectral features in a more intelligent scenario, by weighting the corresponding features for a better representation of the neural components specific to the depth of cognitive processing. By means of this weighted selection, the most representative features for each type of processing would be considered.

Acknowledgement

The work was funded by the sectoral operational programme human resources development 2007-2013 of the ministry of european funds through the financial agreement posdru/159/1.5/S/134398. Further, the research was supported by the EU FP7 programme (FP7/2007-2013, grant agreement no. 611570) and BMBF (grant no. 01GQ0850). The work has received additional support from the Politehnica University of Bucharest through the research grant UPB – GEX / NR. 98/26.09.2016, „Driver's fatigue monitoring through heart rate analysis and eye and facial expression tracking".

R E F E R E N C E S

[1] *A. Bashashati, M. Fatourechi, R. K. Ward and G. E. Birch*, A survey of signal processing algorithms in brain–computer interfaces based on electrical brain signals, J. Neural Eng. 4, 2007, pp. 32–57.

[2] *J. Farquhar and N. J. Hill*, Interactions between Pre-Processing and Classification Methods for Event-Related-Potential Classification: Best-Practice Guidelines for Brain-Computer Interfacing, Neuroinformatics, April 2013, vol. 11, issue 2, pp. 175–192.

[3] *F. Lotte,* A Tutorial on EEG Signal Processing Techniques for Mental State Recognition in Brain-Computer Interfaces, Ed. E. R. Miranda; J. Castet, Guide to Brain-Computer Music Interfacing, Springer, 2014.

[4] *C. Vidaurre, C. Sannelli, W. Samek, S. Dähne and K. R. Müller*, Machine learning methods of the Berlin brain-computer interface Proceedings Volumes (IFAC-PapersOnline), vol. 48, issue 20, 2015, pp. 447-452.

[5] *B. Blankertz, S. Lemm, M. Treder, S. Haufe, K. R. Muller*, "Single-trial analysis and classification of ERP components a tutorial", NeuroImage no. 56(2), pp. 814–825, 2011.

[6] *B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, K. R. Müller*, "Optimizing spatial filters for robust EEG single-trial analysis", IEEE Signal Processing Signal Magazine, no. 25, pp. 41–56, 2008.

[7] *S. Lemm, B. Blankertz, T. Dickhaus and K-R Müller*, "Introduction to machine learning for brain imaging", Neuroimage, no. 56, pp. 387-399, 2011.

[8]. *I. E. Nicolae, L. Acqualagna, and B. Blankertz*, "Tapping Neural Correlates of the Depth of Cognitive Processing for Improving Human Computer Interaction", 4th International Workshop on Symbiotic Interaction, pp. 126-131, 2015.

[9] *M. Uscumlic*, B. *Blankertz*, "Active visual search in non-stationary scenes: coping with temporal variability and uncertainty", J Neural Eng, no. 13, 2016.

[10] *V. V. Nikulin, G. Nolte, and G. Curio*, "A novel method for reliable and fast extraction of neuronal EEG/MEG oscillations on the basis of spatio-spectral decomposition", NeuroImage, no. 55, pp. 1528–1535, 2011.

[11] *I. E. Nicolae, L. Acqualagna, and B. Blankertz*, "Assessing the depth of cognitive processing for user-state adaptation", 2017. (*submitted to Frontiers in Neuroscience*)

[12] *I. E. Nicolae, L. Acqualagna, and B. Blankertz*, "Neural Indicators of the Depth of Cognitive Processing for User-Adaptive Neurotechnological Applications", Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society, pp. 1484-1487, 2015.

[13] *G. Buzsáki*, "Rhythms of the Brain", Oxford University Press, 2006.