HAL
open science

Operationalization of Conceptual Imagery for BCIs

Nataliya Kosmyna, Franck Tarpin-Bernard, Bertrand Rivet

To cite this version:

Nataliya Kosmyna, Franck Tarpin-Bernard, Bertrand Rivet. Operationalization of Conceptual Imagery for BCIs. EUSIPCO 2015 - 23th European Signal Processing Conference, Aug 2015, Nice, France. pp.2726-2730, 10.1109/EUSIPCO.2015.7362880 . hal-01208428

HAL Id: hal-01208428

https://hal.science/hal-01208428

Submitted on 2 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OPERATIONALIZATION OF CONCEPTUAL IMAGERY FOR BCIS

Nataliya KOSMYNA[2], Franck TARPIN-BERNARD[1], Bertrand RIVET[2]

[1] University Grenoble Alpes, France; [2] Grenoble INP, France.

ABSTRACT

We present a Brain Computer Interface (BCI) system in an asynchronous setting that allows classifying objects in their semantic categories (e.g. a hammer is a tool). For training, we use visual cues that are representative of the concepts (e.g. a hammer image for the concept of hammer). We evaluate the system in an offline synchronous setting and in an online asynchronous setting. We consider two scenarios: the first one, where concepts are in close semantic families (10 subjects) and the second where concepts are from distinctly different categories (10 subjects). We find that both have classification accuracies of 70% and above, although more distant conceptual categories lead to 5% more in classification accuracy.

Index Terms— *Brain Computer Interfaces, Conceptual Imagery.*

1. INTRODUCTION

Brain Computer Interfaces (BCIs) are increasingly being used out of the laboratory for interaction and control tasks. The main criterion aside from having high classification accuracy is to propose to users a system that is seamless and minimally noticeable so the user can concentrate on the task rather than the BCI. For control applications, Motor Imagery (imagining moving one's limbs) is a well-studied and adequate BCI paradigm [1]. It is intuitive and naturally corresponds to the semantics of a control task (e.g., turning left or right in a virtual environment). However, when it comes to selection tasks, SSVEP and P300 (repeated flashing of a target) both use visual simulation unrelated to the semantics of the task that can be uncomfortable. One alternative is to explore the detection of conceptual representations and categories in order to build a BCI with a selection paradigm close to the semantics of the task. Conceptual representations, (e.g. a cat is in the broad semantic category of animal) in EEG data have long been studied (e.g. what happens in the brain when a user is stimulated by the concept "cat", seeing a cat, reading the word "cat", hearing it). Recent works have attempted detecting conceptual categories automatically from such signals (EEG, MEG) [2] and found that and that it is possible to classify objects belonging to semantically distinct categories (e.g. cat is an animal, car is a vehicle) as opposed to closely related categories (e.g. a hammer is a tool and a pen is a writing instrument). Simanova et al. [3] go a step further towards building BCI system based on conceptual category detection by studying event-related potential (ERP) activations. They consider textual, visual and auditory cues for concept stimulation. The best performing detection results from visual stimuli. Starting from their findings we intend to build a BCI system that allows detecting the semantic categories of a few concrete concepts (e.g. a lamp, a chair). We thus obtain a BCI paradigm compatible with the semantics of selection tasks. We follow the principles of the theoretical BCI architecture by Kosmyna et al. [4] as well as the critique by Lotte et al. [5] on BCI training protocols to build an asynchronous and incremental BCI system that implements a detection of semantic categories in order to trigger discrete actions that are related to a particular categories of interest (e.g. imagining a lamp so as to turn a light on or off).

We study two settings: one where the concepts are from distinct semantic categories (a goat and a fireplace) and another where the two categories are closer together (a rabbit and a werewolf). We want to answer the following questions:

- (Q1) Is such a system feasible?
- (Q2) How does using closer semantic categories affect the performance of the system?

We will first present the design and architecture of our BCI system. Subsequently, we present the experiments and results along with a discussion and conclude about our findings. Finally we present perspectives for future work.

2. RELATED WORK

Murphy et al. [6] first study the detection of concept categories in EEG and later in MEG [2], [7] using a Common Spatial Patterns (CSP) filter and a Support Vector Machine (SVM) classification using a dense coverage EEG acquisition system.

Simanova et al. make a step further towards building BCI system based on conceptual category detection in their study of resulting event-related potentials (ERP) [3]. They consider textual, visual and auditory cues for concept stimulation. The best performing detection results from visual stimuli.

3. BCI SYSTEM PROPOSED

The rationale behind our BCI system is to make no assumption about the specific BCI paradigm, to require little initial training and to allow for the incremental addition of training trails. We want the user to start using the system right away, even though the performance is potentially sub-optimal and allow for the incremental addition of new training trials if the performance does not satisfy the user

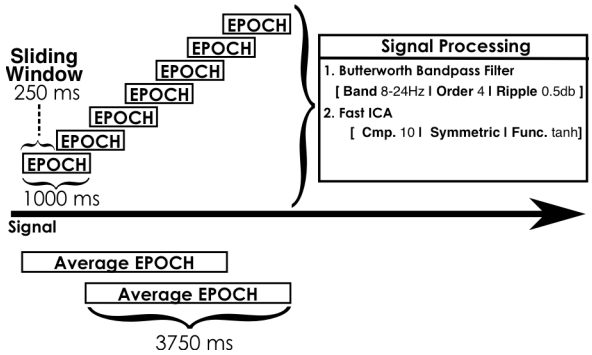

Figure 1. Signal epoch segmentation and averaging in our system.

for the particular task at hand. We follow the architectural principles proposed by Kosmyna et al. [4] that were made following the critique of BCI training protocols by Lotte et al [5]. We use Minimum Distance Classification (MDC), a popular classification approach in pattern recognition [8] and that is compatible with the requirements of the architectural principles we follow and shows good performance for some BCI experiments, especially in an asynchronous setting (e.g. with the Mahalanobis distance) [9]. The best performance is achieved with paradigm-specific distance measure [10]. In our work, we make no assumptions about the nature of the signals and use Independent Component Analysis (ICA) as a preprocessing step with the objective that ICA will project the data in a space closer to the intrinsic topology of the data [11] and make it easier for distance measure to accurately capture relevant differences in EEG patterns. This implies that the number of components computed is lower than the number of sensors.

3.1. Signal Processing & Acquisition

Figure 1 gives a graphical representation of how signals are divided into time 1s epochs over a 250ms sliding window, then filtered, and averaged. Subsequently, signal processing is applied on each epoch:

- The Butterworth filter allows selecting the appropriate frequencies for Conceptual Imagery and discarding unwanted frequencies. We selected a 8-25Hz pass band with a filter of magnitude 4 and with a ripple of 0.5db.
- The FastICA [12] algorithm projects the signal data in a space where data points are maximally independent, essentially separating task related sources from noise sources and other interference. The difference with other ICA algorithms is that FastICA uses a fast algorithm based on fixed-point calculations. We applied the variant of FastICA with symmetric orthogonalization and using a hyperbolic tangent contrast function. We computed 10 components out of 16 sensors. We used a GPL Java implementation of the original algorithm as described by Hyvärinen and Oja [12].
- We then produce average epochs that allow us to smooth the signal and remove some of the variability. The system thus produces an average epoch every second, which is then used for feature extraction and for classification. Thus, the classifier will yield one classification per second. Given that ICA is rather costly to compute, anything less than one second led to sub real-time performance on the machine we performed the processing on (2012 MacBook Air, i7@2.9GHz).

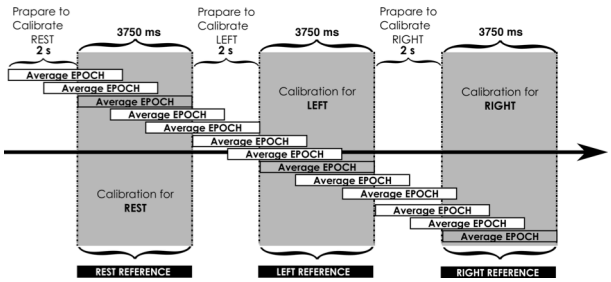

Figure 2. Detailed view of the calibration phase.

Given the sensitivity to noise and to variability in minimum distance based classification, ICA separates noise sources from authentic signals and makes it easier for distance measure to accurately capture relevant differences in brainwave patterns (if we generate less components than sensors).

Our classifier only requires minimal training data to start functioning. Indeed, our aim is to reduce that training time to a single calibration trial per class before the system starts functioning. Subsequently, further training can be incrementally added depending on the initial performance and the satisfaction of the user. For each trial, we capture a reference signal for each of the classes. In the case of our system, we captured full averaged epochs as references. Thus, the calibration for each class lasted 5 seconds and we left 2 seconds of rest between each calibration trial so that users could unwind briefly and prepare for the next trial. Figure 2 graphically shows the process by which we capture the reference signal for a particular BCI task.

3.2. Feature Extraction & Classification

As described in Figure 3, for each classification, we take the current average epoch and compute the distance between this current epoch and the reference signals for each of the class references. We then consider that the class corresponding to the reference epoch that yields the minimum distance is the outcome of the classification. Right after calibration, there is a single reference signal per class, however when feedback is given, more reference

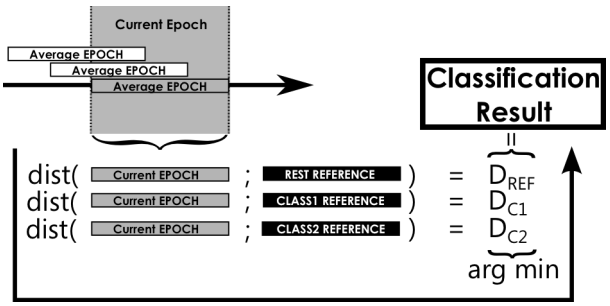

Figure 3. Our classifier.

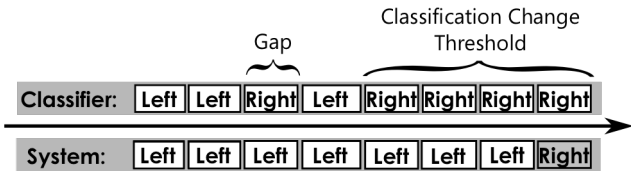

Figure 4. The classification change threshold.

signals are added for each class. When there are several references per class, there will be several distance measurements, in which case the classification will be decided by a majority vote on classification outputs resulting from the individual distance measurements.

Similarly, given that several EEG channels are used, if we use single variable distance measures (as opposed to multivariate measures), we obtain one distance value per channel, which is handled the same way as in the multiple reference setting. We can also consider using several different distance measures at the same time, so long as distances are comparable (same value range and distribution).

In order to minimize the sensitivity of the classifier and rapid successive classification changes, we added a threshold of successive identical classification outcomes before the system classification output changed. As can be seen in Figure 4, initially the classification result is Class 1 (CL.1) twice consecutively and then the classifier gives a Class 2 (CL.2), however the system continues to output Class 1. For the system to change its output to Class 2, we need 4 successive Class 2 classifications from the classifier (with no gaps).

3.3. Equipment and Electrode Placement

For the electrode placement of Conceptual Imagery, we mostly followed the placement proposed by Simanova et al. [3], where 64 electrodes are positioned in an equidistant manner so as to locate and study the underlying phenomena. They make three experiments where visual, textual and auditory stimulation cues are used and find that visual cues lead to the best performance. We focus on visual stimulation and place our electrodes on the activation areas over the visual cortex.

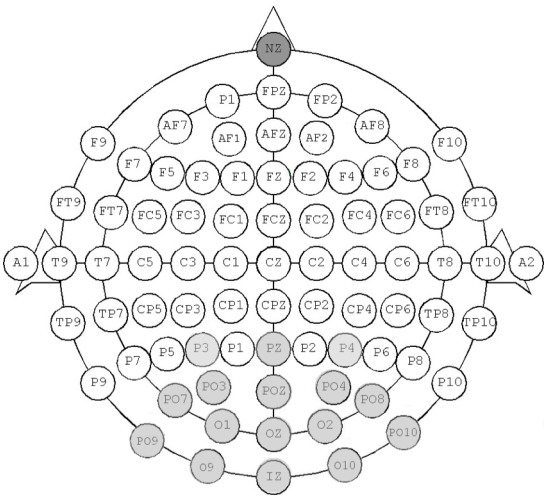

Figure 5. Electrode placement for our BCI according to the extended 10-20 electrode placement system.

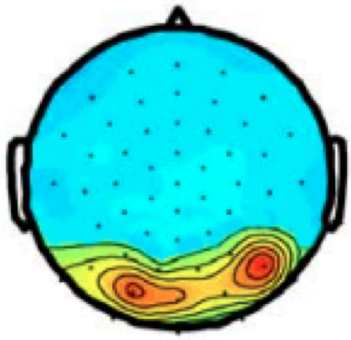

Figure 6. Activation map in the 160-200ms epoch after stimulation.

We used a g.tec USBAmp with 16 channels with g.SAHARA dry electrodes over the following 10-20 electrode positions: P3, P4, Pz, PO9, PO7, PO3, PO4, PO8, PO10, POz, O9, O10, O1, O2, Oz, Iz (Figure 5). Moreover, we plotted an average activation map averaged for one of our subjects in Figure 6 in the 160-200 ms. time interval after stimulation. The activations are indeed identifiable over the visual cortex at similar locations as those identified by Simanova et al., there most likely are activations elsewhere, however we only cover the visual cortex.

4. EXPERIMENTS

In order to answer our questions on whether obtaining a functional BCI system in this context is feasible (Q1) and how using items from closer conceptual categories impact performance (Q2), we made two experiments and for each of the settings we are interested in studying:

- Two concepts from closely related conceptual categories;
- Two concepts from clearly distinct conceptual families.

We first present the two settings and the concepts involved. Then we validate the classification procedure by doing a synchronous training capture of 20 trials for each class and by using 10-fold cross-validation. Finally we evaluate the system in the asynchronous setting, by giving the users objectives in a practical task and counting the number of errors and the time to correct classification in order to evaluate the performance in a real setting. For a two-class classification problem, the system also considers the resting state of the user that should include all activity except the target activity. This is essential for an asynchronous setting as the classifier is constantly active, even when the user is at rest, in which case the classifier should recognize it [13].

4.1. Similar Conceptual Families

To study similar conceptual family detection, we developed a simple game where users played with the little riding hood and had to go home, but at each step there could be a rabbit menaced of being eaten by the big bad

(Were)wolf or the player could be menaced of being eating by the wolf directly. The user had to imagine the concept of a rabbit to save the rabbit and dash forward (in the first situation) or to imagine the wolf to kill it (in the second situation). A rabbit is an animal but a werewolf is not quite an animal, although close to one. The question for this setting is how this closeness affects the performance of the BCI. The user had to chose between the rabbit or the werewolf at each turn over 16 turns.

4.2. Different Conceptual Families

In the second setting we devised a similar game with a different objective at each turn. A herder's goats have run away far and wide. The herder wants to get the goats back and goes on a journey. At each turn, the herder either hears a goat and the player must imagine a goat to catch the goat in the game or if night falls (the sky darkens), the player must imagine a fire in order to set camp for the night. The herder finished the journey in 8 days and nights (16 sessions), however for the experiments the order was randomized. This setting serves to answer the first experimental question of whether BCI performance will be acceptable in out restricted setting compared to Simanova et al.'s study.

5. RESULTS & ANALYSIS

5.1. Synchronous Offline Validation

We performed a simple validation that consisted classifying a set of unlabeled signals and comparing them to a set of reference signals for each class (including the rest class) using 10-fold cross-validation. We considered two distance measures in order to see what was the one (if any) leading to the best performance. The analysis was done over the signals of two subjects (distinct from subjects of the asynchronous experiments in the next session) captured over the course of 20 trials for each class. We used both the Mahalanobis distance (parametric – assumes a normal distribution, multivariate – one measure for all channels at once) and the spearman rank correlation coefficient (non-parametric, single variate – distance averaged over channels to obtain a single measure). We can now look at the cross-validation results from the off-line analysis to validate the BCI system depending on the distance measure used in Table 1. In bold is the best result for each measure and each subject. The Mahalanobis distance performed better with a significant difference (ANOVA + Tukey HSD post-hoc with p-value correction, with $p<0.01$) from the spearman difference except for subject two in the context of different conceptual families. The difference of performance between subjects was not significant. Moreover the results were not random following the empirical study by Müller-Putz et al. [14].

Distance Measure	**Concept Distinction Type**			
	Close conceptual family		***Different conceptual families***	
	Subject 1	**Subject 2**	**Subject 1**	**Subject 2**
Mahalanobis	**71.34%** (1.23%)	**71.45%** (1.14%)	**74.23%** (0.87%)	**73.96%** (0.75%)
Spearman	69.32% (2.03%)	68.94% (1.87%)	72.45% (1.22%)	72.65% (1.43%)

Table 1. Synchronous Offline Analysis in terms of 10-fold cross-validated Classification Accuracy (STD. Dev. in parenthesis) for both subjects and for the spearman rank correlation and Mahalanobis distances.

Close conceptual family		***Different conceptual families***	
Accuracy	**Time**	**Accuracy**	**Time**
70.12% (0.65%)	2512ms (234ms)	**75.62% (0.45%)**	**1245ms (143ms)**

Table 2. Asynchronous evaluation. Classification Accuracy (%) and average classification time (MS), standard deviation in parenthesis.

5.2. Asynchronous online evaluation

For the asynchronous online evaluation, we had 10 users for each setting (10 subjects for close categories and 10 other subjects for distant categories) and had the users perform the respective task. Users were between 23 and 45 years old and all had prior BCI experience. In the first setting the user had to chose between the rabbit or the werewolf at each turn over 16 turns without a predefined duration or stimulus (only one was correct, the order was randomized). The turn ended as soon as the classification output was different than the rest state. Similarly, in the second setting users had to cross 16 segments of a game. Similarly the user got out of the current segment only when the classification output was different from rest. For both situations we measured the average classification accuracy and the average classification time (time between the start of the turn or segment and the moment the classification output becomes different from rest). Table 2 shows the results. The difference in the means of the classification accuracy and classification time between the two settings is significant with $p<0.01$, using a student t-test (the assumption of equality of variances holds). Like in the synchronous setting, both setting lead to a usable classification performance for the BCI, however with close conceptual categories is 5% lower than with distinct conceptual categories (70.12% vs. 75.62%). The difference in classification time is even more marked as it takes about half as long to perform a classification with distinct conceptual categories rather than with close conceptual categories (1245ms against 2512ms).

5.3. Discussion

From both the synchronous validation setting and the asynchronous evaluation setting, we can answer positively to Q1 about the fact that we can indeed build a functional

BCI system using conceptual category classification. Even though we only concentrate on one stimulus modality and area in the brain. Arguably covering a broader area would allow us to capture higher level processing of the conceptual information and allow for the better detection of closer categories. Indeed, while both systems are usable, the 5% difference in classification accuracy and the two-fold increase in classification time have a potential effect on the usability of the system (Q2) for the end users.

5.4. User opinions

The subjects of the experiment were very interested by the conceptual imagery BCI paradigm as they saw great potential for an engaging interaction in many contexts. However, for close conceptual families they find the delay it took for a successful classification to occur rather frustrating as opposed to the classification time for distinct conceptual families that was deemed tolerable.

6. CONCLUSION

We build an operational BCI system in an asynchronous setting that allows classifying objects in their semantic categories based on visual cues. We evaluate the system in an offline synchronous setting and an online asynchronous setting in two situations: two concepts in families that are close and two concepts from distinctly different categories. We find that both have classification accuracies of 70% and above, although more distant conceptual categories lead to 5% more in classification accuracy in the online setting. For future work, the inclusion of a broader coverage of the brain would potentially allow for finer grained distinction between categories. Moreover given that we only used two classes in this experiment, it is necessary to extend and evaluate this work to more classes in order to see how the performance and training time required scale up.

REFERENCES

[1] E. Thomas, M. Dyson, and M. Clerc, "An analysis of performance evaluation for motor-imagery based BCI.," *J. Neural Eng.*, vol. 10, no. 3, p. 031001, Jun. 2013.

[2] B. Murphy, M. Poesio, F. Bovolo, L. Bruzzone, M. Dalponte, and H. Lakany, "EEG decoding of semantic category reveals distributed representations for single concepts.," *Brain Lang.*, vol. 117, no. 1, pp. 12–22, 2011.

[3] I. Simanova, M. van Gerven, R. Oostenveld, and P. Hagoort, "Identifying object categories from event-related EEG: toward decoding of conceptual representations.," *PLoS One*, vol. 5, no. 12, p. e14465, Jan. 2010.

[4] N. Kos'myna, F. Tarpin-Bernard, and B. Rivet, "Towards a general architecture for a co-learning of brain computer interfaces," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2013, pp. 1054–1057.

[5] F. Lotte, F. Larrue, and C. Mühl, "Flaws in current human training protocols for spontaneous Brain-Computer Interfaces: lessons learned from instructional design.," *Front. Hum. Neurosci.*, vol. 7, no. September, p. 568, Jan. 2013.

[6] B. Murphy and M. Dalponte, "Distinguishing concept categories from single-trial electrophysiological activity," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, 2008, pp. 403–408.

[7] B. Murphy and M. Poesio, "Detecting semantic category in simultaneous EEG/MEG recordings," in *Proceedings of the naacl hlt 2010 first workshop on computational neurolinguistics*, 2010, no. June, pp. 36–44.

[8] R. Duda, P. Hart, and D. Stok, *Pattern Recognition, second edition*. WILEY-INTERSCIENCE, 2001.

[9] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interfaces.," *J. Neural Eng.*, vol. 4, no. 2, pp. R1–R13, Jun. 2007.

[10] M. Congedo, A. Barachant, and A. Andreev, "A New Generation of Brain-Computer Interface Based on Riemannian Geometry," *arXiv Prepr. arXiv1310.8115*, vol. 33, no. 0, 2013.

[11] M. Naeem, C. Brunner, R. Leeb, B. Graimann, and G. Pfurtscheller, "Seperability of four-class motor imagery data using independent components analysis.," *J. Neural Eng.*, vol. 3, no. 3, pp. 208–16, Sep. 2006.

[12] A. Hyvärinen and E. Oja, "A Fast Fixed-Point Algorithm for Independent Component Analysis," *Neural Comput.*, vol. 9, pp. 1483–1492, 1997.

[13] S. Mason, J. Kronegg, J. Huggins, M. Fatourechi, and A. Schlögl, "Evaluating the Performance of Self - Paced Brain-Computer Interface Technology," *Neil Sq. Soc.*, 2006.

[14] G. R. Müller-Putz and R. Scherer, "Better than random? A closer look on BCI results," *Int. J. Bioelectromagn.*, vol. 10, no. 1, pp. 52–55, 2008.