DOI:10.3217/978-3-85125-260-6-62

Iterative EEG-Based Natural Image Search Under RSVP

M. Ušćumlić, R. Chavarriaga, J. del R. Millán

Defitech Chair in Non-Invasive Brain-Machine Interface, Center for Neuroprosthetics, EPFL, Lausanne, Switzerland

Correspondence: STI-CNBI École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland.
E-mail: marija.uscumlic@epfl.ch

Abstract. This work extends previous studies on using EEG decoding for automatic image retrieval. We propose an iterative way to integrate the information obtained from the EEG decoding and image processing methods. In the light of real-world BCI applications, we demonstrated that a limited number of EEG channels provide sufficient information about the subject's preference to be exploited in image retrieval by the proposed synergistic scenario. Furthermore, to meet a more realistic scenario we used natural images (i.e., images of objects in their natural environment).

Keywords: EEG, Single-Trial Classification, RSVP, Image Retrieval, BCI

1. Introduction

Humans ability to process visual information outperforms state-of-the art computer methods. For this reason, analysis of EEG responses to visual stimuli has been proposed as a complement to image recognition systems. In particular, using the rapid serial visual presentation (RSVP) protocol. In this scenario, the presented images are labeled based on the EEG activity as target/non-target. Then, the decoded labels are propagated to unseen images based on similarity and data mining methods [Pohlmeyer et al., 2011].

We propose an alternative iterative scenario for coupling the EEG decoding with automatic image labeling. An iteration consists of assigning the EEG-based labels to the presented images (i.e., RSVP sequence) and their propagation to the unseen images. This yields a set of probabilistic labels based on both brain signals and image features. Then, we fuse the labels obtained at each iteration before ranking the whole database and retrieve that target images.

2. Material and Methods

2.1. Experimental Setting

Subjects ($N = 15$) were presented with sequences of natural images at a rate of 4 Hz. They were instructed to count images of a specified object. The experiment consisted of two phases: training and testing. Different sets of images from Corel database were used in the training (1600 images) and testing phases (1382 images). Four search tasks (Elephant, Car, Lion and Butterfly) were given in the training phase, and three search tasks (Eagle, Tiger and Train) in the testing phase. In the training sequences 10 % of images were the targets. The testing phase consisted of four iterations (200 images per iteration). In the initial iteration, a sequence of images was presented (10 % of them targets). The elicited EEG response to each image was decoded to obtain labels for the presented images (target/non-target). This information was used to label the remaining images in the database and obtain the image sequence that was shown in the next sequences.

EEG data were recorded with a 64-channel BioSemi ActiveTwo system, at a sampling frequency of 2048 Hz. The EEG signals were bandpass filtered [1 10 Hz] and downsampled to 32 Hz. The EEG signals were re-referenced by common average reference (CAR) based on 41 electrodes (the peripheral electrodes were excluded).

2.2. EEG-based Image Labeling

The EEG signals from the training phase are used to train a Gaussian classifier (target vs. non-target trials) [Millán et al., 2004], using four prototypes per class. The feature vector is obtained by concatenating samples in the interval from 200 ms to 700 ms after stimulus onset of a subset of 8 channels: Pz, PO3, POz, CPz, Cz, PO4, C3, C4. The feature dimensionality is reduced, keeping only the features with high discriminant power (DP) [Galán et al., 2007].

2.3. Automatic Image Labeling

This step propagates the labels obtained from the EEG decoding to the remaining (unseen) images in the database. We used a semi-supervised approach for automatic image labeling, exploiting a visual similarity graph of the images in database [Yang et al., 2006]. Each node in the graph represents an image, while its state is the probability that the

 DOI:10.3217/978-3-85125-260-6-62

image belongs to the target class. Every node is connected to ten neighboring nodes where five of them have been assigned a label based on the EEG. In turn, arcs represent the probability that the connected nodes belong to the same class taking into account their similarity in the image feature space. The images in our database are indexed in two feature spaces: edge histograms and the colored pattern appearance model [Qiu, 2004].

2.4. Iterative Coupling

The EEG-based labels (target/no target) obtained in a given iteration are used for the automatic labeling, i.e., they are used as initial labels in the graph and then propagated to the entire database. Then, a new RSVP sequence for the next iteration is generated from the 200 top ranked images based on their labels. By repeating these steps we are accumulating evidence of the true labels for each image in the database. After the final iteration, the label probabilities for each image are averaged across iterations to obtain the final labeling.

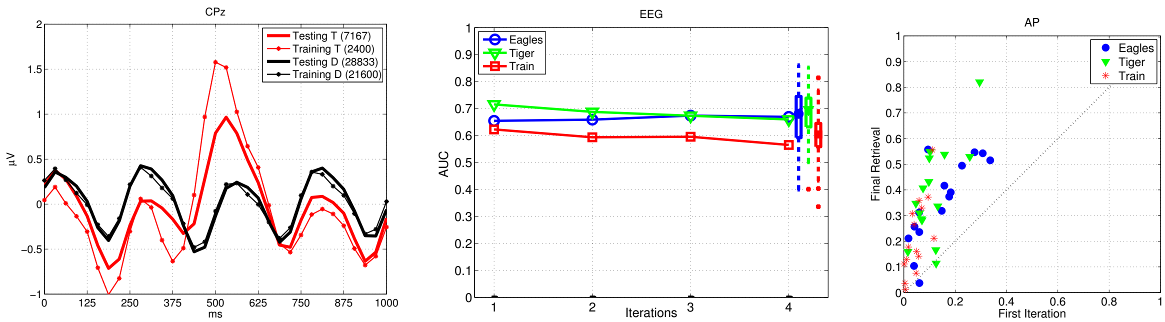

Figure 1: *(a) Grand Average ERP at CPz. T: target; D:non-target; (No of trials). (b) EEG classification performance (AUC averaged across subjects). (c) AP of the top-ranked 200 images across subjects: 4th vs 1st iteration. Colors indicate search tasks.*

2.5. Results and Discussion

ERPs over centroparietal electrodes exhibit a positive peak at about 500 ms after target images are presented (Fig. 1a). The single-trial EEG classification performance, in terms of the area under the ROC curve (AUC) is shown in Fig. 1b. Performance exceeds chance level in all search tasks, although results for task Train were significantly lower (Friedman, $p < 0.05$). This task-dependent differences in EEG performance were consistent with response times (RT) measured in a separate behavioral RSVP protocol using the same images. Accordingly, the longest median RT was found for the class Train, indicating its lower discriminability.

We evaluate the retrieval performance in terms of the average precision (AP) on the top-ranked 200 images (i.e. the mean of the precision scores over these images). A comparison between the retrieval at the first and the last iterations (c.f. Fig. 1c) clearly shows that the iterative coupling results in significant improvement with respect to the labels obtained in the initial iteration. Thus making it less sensitive to EEG labeling errors. These results demonstrate that the limited set of the EEG channels provides sufficient information to make the EEG-based image retrieval operational in the iterative coupling scenario.

Acknowledgments

This study is supported by the Swiss National Science Foundation (grant 200021-120293).

References

Galán, F., Ferrez, P. W., Oliva, F., Guàrdia, J., and Millán, J. d. R. (2007). Feature extraction for multi-class BCI using canonical variates analysis. In *IEEE International Symposium on Intelligent Signal Processing*, pages 1–6.

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

Pohlmeyer, E. A., Wang, J., Jangraw, D. C., Lou, B., Chang, S. F., and Sajda, P. (2011). Closing the loop in cortically-coupled computer vision: A brain-computer interface for searching image databases. *J Neural Eng*, 8(3).

Qiu, G. (2004). Embedded colour image coding for content-based retrieval. *J Vis Commun Image R*, 15:507–521.

Yang, M., Guan, J., Qiu, G., and Lam, K. (2006). Semi-supervised learning based on bayesian networks and optimization for interactive image retrieval. In *British Machine Vision Conference*, page 969.