IMPLICIT RETRIEVAL OF SALIENT IMAGES USING BRAIN COMPUTER INTERFACE

Ashkan Yazdani[1], Jean-Marc Vesin[2], Dario Izzo[3], Christos Ampatzis[3], Touradj Ebrahimi[1]

[1] Ecole Polytechnique Fédérale de Lausanne (EPFL), Institute of Electrical Engineering, Multimedia Signal Processing Group

[2] Ecole Polytechnique Fédérale de Lausanne (EPFL), Institute of Electrical Engineering, Applied Signal Processing Group

[3] European Space and Technology Research Center (ESTEC), Advanced Concepts Team

ABSTRACT

Space missions are often equipped with several high definition sensors that can autonomously collect a potentially enormous amount of data. The bottleneck in retrieving these often precious datasets is the onboard data storing capability and the communication bandwidth, which limit the amount of data that can be sent back to Earth. In this paper, we propose a method based on the analysis of brain electrical activity to identify the scientific interest of experts towards a given image in a large set of images. Such a method can be used to efficiently create an abundant training set (images and whether they are scientifically interesting) with a considerably faster image presentation rate that can go beyond expert consciousness, with less interrogation time for experts and relatively high performance.

Index Terms— EEG, BCI, Signal Processing, Machine Learning, Implicit Retrieval

1. INTRODUCTION

Autonomous decision making systems are becoming increasingly demanded for various purposes. Several research groups across the world are conducting extensive research to increase the performance and simultaneously reduce the cost and risk of the decision in such systems for miscellaneous applications such as space exploration, medicine and militaryapplications. Space exploration is by its very nature an expensive and risky endeavor. Various factors such as stringent communications constraints (limited communication windows, long communication latencies, and limited bandwidth), limited access and availability of operators, limited crew availability, and system complexity restrain direct human oversight of many functions [1]. Therefore, autonomy in space research is far more than just a convenience and is critical to the success of the mission in some situations. Examples of some previous works in this context are [2] and [3]. In [4], an on-board autonomous system was developed and downloaded on Spirit and Opportunity to detect dust-devils and clouds in the Martian landscape. Another prevailing interest in space research is to assess the possibility of elaborating an intelligent module and incorporating it into space exploration robots in extra orbital missions to enable them to search in the immense image datasets that they autonomously collect using several high-definition sensors and select only the scientifically interesting images. Consequently, these robots will be able to discard the rest of images and to transmit only the selected images back to Earth. Such a module can also be used

for automatic detection of anomalies in medical images, interesting military images, etc. in large image sets.

A key point, at the center of current technological developments, is the design of algorithms able to classify sensor readings (e.g. images) according to their degree of scientific interest. The main difficulty lies in the definition of what is scientifically interesting. Machine learning algorithms could be trained directly to classify what is scientifically interesting and what is not, without further information about these two very broad classes. This could potentially allow for broader and more fuzzy classification borders, and could result in algorithms able to return not only the strictly defined and expected, but also a set of images with potentially unexpected, but relevant properties. The challenge when following this approach becomes how to then create a training set for a classifier.

One option is to resort to what is typically referred to as the interviewing or interrogation technique. Expert scientists would be interviewed on a particular set of pictures, being asked to simply classify or rank them; subsequently a computer would be trained to have a similar response to the one of the interviewed scientist. In this way, the computer has to automatically extract the relevant features that guided the expert's decision-making and learn to use them in such a way so to mirror the expert's classification.

Despite the simplicity of this methodology, there are various drawbacks involved. For example, it requires the scientists to undergo long and time-consuming sessions of image classification that may prove to be particularly tiring and cumbersome, which in turn can result in the acquisition of a noisy training set. Moreover, this approach is subject to the fuzziness of the scientist's reasoning when placing a highly cognitive judgment upon each picture. In other words, the scientist will repeatedly consciously filter the image, eventually merging even contradictory verdicts to one binary classification or a ranking. In this work, we propose an alternative approach to creating such a training set for a classifier; in particular, the information about the expert's classification is extracted directly from the classification of his/her brainwaves.

It is well known from neurophysiological studies that when subjects look at images which arouse mental response such as surprise, anticipation, and etc., their parietal cortex is excited in a very characteristic way: a synchronized peak in the global electrical activity of large groups of neurons in the parietal area arises after approximately 300 ms after the stimulus (image) presentation. This electrical activity can be recorded with an Electro-Encephalography (EEG) instrument as an electric positive potential wave and is commonly referred to as P300 [5].

We propose to extract the picture rating information using the EEG signal recorded while the expert is presented with the pictures in a Rapid Serial Visual Presentation (RSVP) experiment. Our set-

up is inspired by a related work performed by Gerson et al. [6] who presented an original experiment where a simple image ranking task is performed by ranking images according to the amplitude of the P300 brainwave recorded during a RSVP experiment based on the oddball paradigm. The oddball paradigm refers to experimental setups in which a target stimulus is presented among more frequent background stimuli. The results of this experiment, and the vast pre-existing literature available on the detection and use of the P300 wave for different applications such as Brain Computer Interface (BCI) [7], suggests the potentiality of processing EEG signals recorded during a RSVP experiment using machine learning techniques to extract a classification of the images mirroring the classification the subject made when presented with the images. In other words, the EEG signal could be used to define image classes rather then having the scientist analyze and explicitly perform such a classification.

The rest of this paper is organized as follows. In Sections 2 and 3, the experiments performed and the signal processing and classification algorithms used in this study are described . A discussion of results follows in Section 4 and finally Section 5 concludes the paper.

2. EXPERIMENTS

The aim of the designed experiments were to answer the following questions.

1. Can the P300 signal be reliably detected with the experimental set-up used and with the available tools?

2. How is P300 detection reliability affected by the the image presentation rate (i.e. the number of images presented per second)?

In order to fulfill these objectives, the classical oddball paradigm was used throughout the experiments. Visual stimuli consisted of a set of 3204 images of gray stones illuminated with a uniform ambient light. 25 of those images contained in addition to the stones, a sand model of a spacecraft, thus constituting the oddball images. The spacecraft position was different in each of these images but the object itself was clearly visible in all cases. Examples of background and oddball images for this experiments are displayed in figure 1.
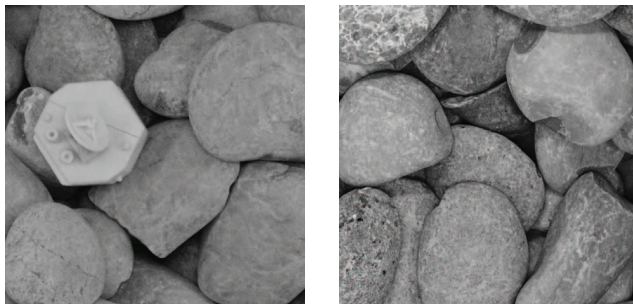

Fig. 1. Examples of an oddball (left) and background (right) images used in the experiments.

To answer the above-mentioned questions, five experiments were designed and performed. Each experiment involved the presentation of five image sequences to subjects. Four subjects (all male, and right-handed) were asked to participate in our experiments and each of them completed two recording sessions. The second session was performed on another day but for all subjects

No. of subjects	No. of sequences	Images in seq.	Oddballs in seq.	Repetitions	IDP/IIP (ms)	T (s)
4	5	40	4	2	500/500	40
4	5	67	7	2	300/300	40
4	5	133	13	2	150/150	40
4	5	200	20	2	100/100	40
4	5	400	40	2	50/50	40

Table 1. Parameters of the different experiments

the time between the first and the last session was less than one week. Before these sessions, the subjects were instructed to watch the image sequences and silently count the number of images containing the spacecraft model and were made familiar with examples of an oddball and non-oddball image. It is worth mentioning that the counting task was imposed merely to keep the subject concentrated during the test. Furthermore, at the end of each image sequence, the subject was asked to report the result of counting and this result was used to monitor the performance of the subject. During each session, the actual image sequences were presented with the EEG signals being recorded at the same time, always preceded by a countdown screen of duration five seconds that allowed the subjects to prepare for the experiment, reducing the surprise effect of the sequence start.

In each experiment, images were presented to the subject with different predefined Image Display Period (IDP) values, followed by a period of neutral background appearance or Inter Image Period (IIP), resulting in a different image per second presentation rates. The parameters of these experiments are presented in Table 1.

Image sequences of different lengths were presented to the subjects with increasing image presentation rate. The number of images was adjusted to the change in presentation rate, so that the total duration of one sequence remained equal to 40 seconds across all experiments. The number of oddball images present in the sequence was adjusted accordingly, so that the ratio of the number of oddball images, randomly placed in the sequences, to the number of non-oddball images was kept at the same level (10%).

3. SIGNAL PROCESSING AND CLASSIFICATION

In each experiment, the image sequences were presented to the subjects and the exact system clock of each image presentation onset was stored. The EEG signal was recorded synchronously with image presentation at 2048 Hz sampling rate from 32 active electrodes placed at the standard positions of the 10-20 international system. A Biosemi Active Two amplifier was used for amplification and 24-bit analog to digital conversion. After pre-amplification and digitalization of the EEG signal, a block of preprocessing filters was applied to the raw signals to remove the various artifacts. Wavelet decomposition was then applied to break down the signal into different frequency components and for further multiresolution analysis. In the next step, Features are extracted from the signal and finally the extracted feature vectors were classified into either target (salient image) or non-target (normal image) classes. Preprocessing methods were applied in the order stated below:

1. Referencing: The average signal from all electrodes was used for referencing.

2. Filtering: A 12th order forward-backward Butterworth bandpass filter with zero phase shift and the cut-off frequencies of 1 Hz and 95 Hz was used to remove all high frequency

non-EEG artifacts as well as low frequency drifts from the recorded signal.

3. Downsampling: In order to reduce the number of samples, the EEG was downsampled from 2048 Hz to 256 Hz using an eight-order Lowpass Chebyshev Type I filter.

After this preprocessing phase, Discrete Wavelet Transform (DWT) was used to decompose the EEG signal into its subbands. Wavelet transform is known to be effective for representing various aspects of signals such as trends, discontinuities, and repeated patterns when other signal processing approaches fail or are not as effective. The EEG signal contains non-stationary components and it has been shown that wavelets clearly outperform Fourier transform based methods for feature extraction from the EEG signals [8] .

The Daubechies family of wavelets is one of the most commonly used orthogonal wavelets satisfying the admissibility conditions, thus allowing reconstruction of the original signal from the wavelet coefficients. Daubechies wavelets of different orders were investigated for the analysis of epileptic EEG records. Daubechies of order 4 and higher were found to be the most appropriate for analysis of EEG data. In this study, Daubechies order 6 was used for decomposition of EEG signals into different bands. Figure 2 illustrates a typical EEG signal used in this study and its decomposition using DWT. It has been shown in [8] that only the delta and theta frequencies play a major role in the P300 response, however it was discussed that alpha frequencies have also minor roles. Therefore, these three subbands (the last four rows in Figure 2, which correspond to 1-2Hz, 2-4Hz, 4-8 Hz, and 8-16 Hz) were summed together to reconstruct the P300 signal.
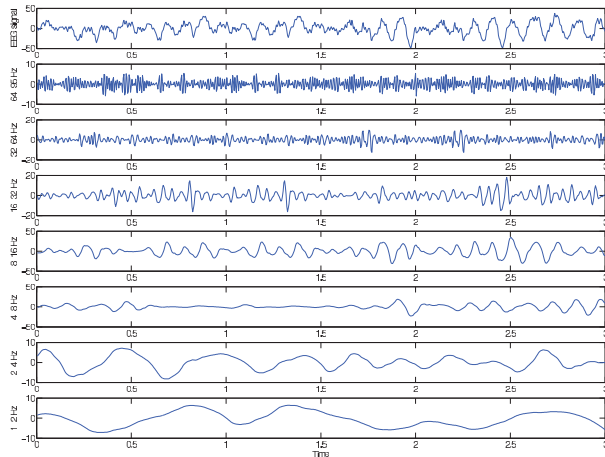

Fig. 2. Three seconds EEG signal and its decomposition

In the next step, single trials of duration 1000 ms were extracted from the data. Single trials started at stimulus onset (provided by the image presentation software), i.e. at the beginning of the intensification of first the image in the image sequence, and ended 1000 ms after stimulus onset. This, in faster experiments, there is some overlap between the single trials of successive images. For instance, for IDP=IIP=150 ms , the last 700 ms of each trial were overlapping with the first 700 ms of the following trial. The EEG signal was again downsampled from 256 Hz to 32 Hz using an eight-order Lowpass Chebyshev Type I filter. Therefore, the decimated signal contains frequency components of 1-12.8 Hz and each single trial contains 32 samples. 16 electrodes (central, parietal and occipital)were se-

lected and the signals of these electrodes were concatenated to form a feature vector corresponding to that single trial.

Eyeblinks, eye movement, muscle activity, or subject movement can cause large amplitude outliers in the EEG. To reduce the effects of such outliers, the data from each electrode were windsorized. For the samples from each electrode the 10th percentile and the 90th percentile were computed. Amplitude values lying below the 10th percentile or above the 90th percentile were then replaced by the 10th percentile or the 90th percentile, respectively.

A Support Vector Machine (SVM) classifier with radial basis function kernels was used for classification of single trials. To perform the classification using this method, five-fold cross validation was performed and it was repeated 15 times. In each training run, the extracted feature vectors was first preprocessed using Principal Component Analysis (PCA) and the dimension of the feature vectors was reduced from 512 to 288. The Mixing matrix W was used to be applied later for dimension reduction on test data. In the next step, the training and test data were normalized to have the minimum and maximum values of zero and one, respectively. Finally, the training data was used to train the SVM classifier. To this end, a grid search for parameter c, was performed. However, the value of parameter γ was fixed based on the information of the training data as follows. The radial basis kernel (also called gaussian kernel) can be expresses as:

$$k(\mathbf{x_i}, \mathbf{x_j}) = \exp(-\gamma \|\mathbf{x_i} - \mathbf{x_j}\|^2), \qquad for\ \gamma > 0 \qquad (1)$$

Therefore, to determine the value of γ we computed the average distance between two feature vectors of the same class in the training set (d) and we set the gamma value as:

$$\gamma = \frac{1}{d} \qquad (2)$$

4. RESULT

In this section, the results of different experiments described in section 2 are discussed. To begin the analysis of signals corresponding to target and non-target images, the target and non-target signals were averaged separately. Averaged signals, showed that all the parietal, parietal-occipital and central electrodes contain the P300 pattern. Moreover, the P300 can be also observed in occipital channels.

The P300 amplitude is usually larger in parietal regions (largest in PZ electrode) compared to other parts of the scalp. Figure 3 illustrates the amplitude of P300 in PZ in various experiments. In this figure, the red and blue signals represent the averaged targets and averaged non-targets, respectively. It can be seen that there is a clear P300 peak in the averaged target signals. Furthermore, the amplitude of P300 clearly decreases as the speed of image presentation increases. In the first experiment, the largest amplitude of P300 equals $14.03\mu v$ and as the image presentation rate increases, this value degrades to $9.57\mu v$, $8.93\ \mu v$, $8.38\ \mu v$, and finally $5.66\ \mu v$ in the fastest experiment. The reason for this might be twofold. First, as the speed of image presentation increases, it becomes harder for the subjects to detect the target images in the sequence and second, as the image presentation rate increases, the evoked potentials interfere with each other and the measured EEG signal is superposition of different evoked potentials.

After the analysis of averaged signals, single trial analysis was performed using the methodology introduced in Section 3, and the computed feature vectors were classified using the SVM classifier with radial basis kernel.
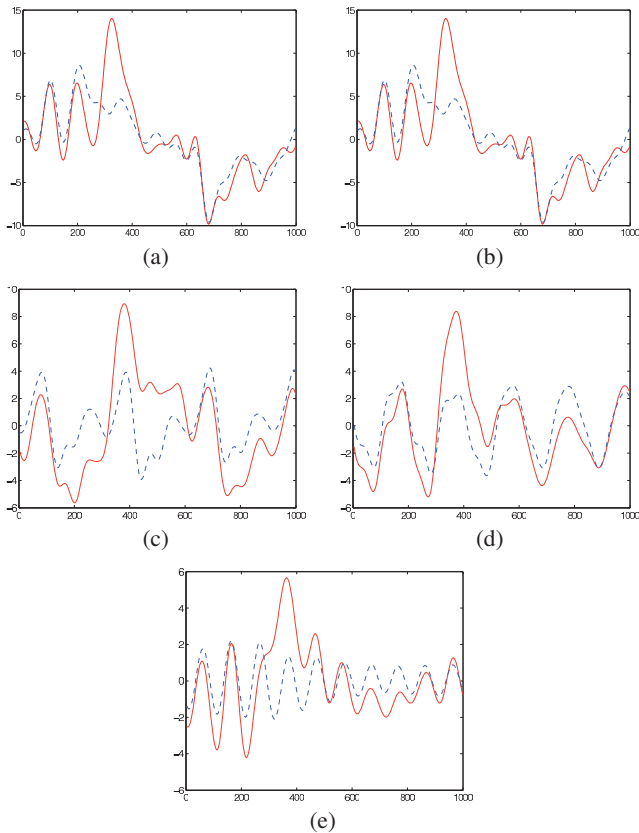
(a)

(b)

(c)

(d)

(e)

Fig. 3. Averaged signals of PZ electrodes for different image processing rates (IIP=IDP= 500 ms (a) to IIP=IDP= 50 ms (e)). X axis corresponds to the time after the stimulus onset in ms and Y axis displays the amplitude of the P300 signal in μv.

In order to evaluate the performance of the system, the precision and recall values were computed for each subject and each experiment. In the scenario used in this study, precision is defined as the number of retrieved salient images divided by the total number of images retrieved using the designed BCI system, and recall is defined as the number of salient images retrieved using the designed BCI system divided by the total number of existing salient images in the image sequences (which should have been retrieved). Usually, precision and recall scores are not discussed in isolation. Instead, either values for one measure are compared for a fixed level at the other measure (e.g. precision at a recall level of 0.75) or both are combined into a single measure, such as the F-measure, which is the weighted harmonic mean of precision and recall.

Figure 4 illustrates the result of the classification. As it can be seen in this figure, relatively high F-measure was obtained in the experiments. Furthermore, F-measure degrades dramatically as the image presentation rate increases.

5. CONCLUSION

In this work, an efficient P300-based BCI system for classification of brainwaves associated to scientific interest in image stimuli was presented. It was shown that relatively high classification accuracies can be obtained for most of the users of this system. It has
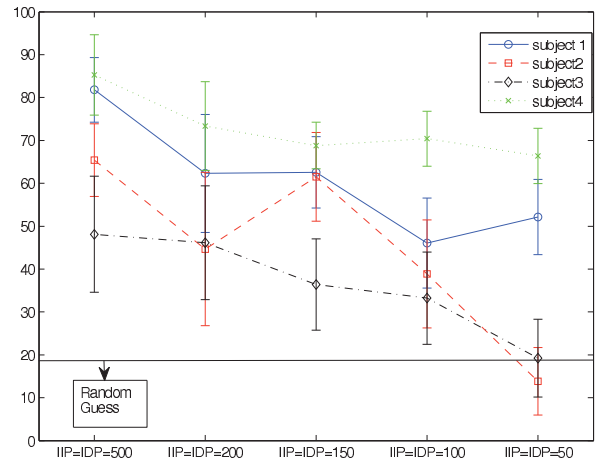

Fig. 4. Comparison of changes in F-measure values between all subjects with respect to speed using SVM classifier.

been observed that increasing the speed of image presentation, will decrease the classification accuracy. However with an image presentation frequency of 3.33 Hz, it is still possible to have relatively good classification results for most of the subjects.

6. REFERENCES

[1] A. Jonsson, R.A. Morris, and L. Pedersen, "Autonomy in Space: Current Capabilities and Future Challenge," *AI magazine*, vol. 28, no. 4, pp. 27, 2007.

[2] S. Chien, R. Sherwood, D. Tran, B. Cichy, G. Rabideau, R. Castano, A. Davies, D. Mandl, S. Frye, B. Trout, et al., "Using autonomy flight software to improve science return on Earth Observing One," *Journal of Aerospace Computing, Information, and Communication*, vol. 2, no. 4, pp. 196–216, 2005.

[3] L. Mandrake, K.L. Wagstaff, D. Gleeson, U. Rebbapragada, D. Tran, R. Castaño, S. Chien, and R.T. Pappalardo, "Onboard Detection of Natural Sulfur on a Glacier via a SVM and Hyperion Data," *In IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2009.

[4] A. Castano, A. Fukunaga, J. Biesiadecki, L. Neakrase, P. Whelley, R. Greeley, M. Lemmon, R. Castano, and S. Chien, "Automatic detection of dust devils and clouds on Mars," *Machine Vision and Applications*, vol. 19, no. 5, pp. 467–482, 2008.

[5] T.W. Picton, "The P300 wave of the human event-related potential," *Journal of clinical neurophysiology*, vol. 9, no. 4, pp. 456, 1992.

[6] A.D. Gerson, L.C. Parra, and P. Sajda, "Cortically coupled computer vision for rapid image search," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 14, no. 2, pp. 174, 2006.

[7] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, and T.M. Vaughan, "Brain–computer interfaces for communication and control," *Clinical neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.

[8] V. Kolev, T. Demiralp, J. Yordanova, A. Ademoglu, and "U. Isoglu-Alkaç, "Time-frequency analysis reveals multiple functional components during oddball P300," *NeuroReport*, vol. 8, no. 8, pp. 2061, 1997.