CURIOSITY CLONING: NEURAL ANALYSIS OF SCIENTIFIC KNOWLEDGE

Dario Izzo[1], **Luca Rossini**[1], **Marek Rucinski**[1], **Christos Ampatzis**[1], **Graham Healy**[2], **Peter Wilkins**[2], **Alan Smeaton**[2], **Ashkan Yazdani**[3], **and Touradj Ebrahimi**[3]

[1]*Advanced Concepts Team, European Space and Technology Research Center (ESTEC)*
[2]*CLARITY: Centre for Sensor Web Technologies, Dublin City University*
[3]*Multimedia Signal Processing Group, Institute of Electrical Engineering, Ecole Polytechnique Fdrale de Lausanne (EPFL)*

ABSTRACT

Event-related potentials (ERPs) are indicators of brain activity related to cognitive processes. They can be detected from EEG signals and thus constitute an attractive non-invasive option to study cognitive information processing. The P300 wave is probably the most celebrated example of an event-related potential and it is classically studied in connection to the odd-ball paradigm experimental protocol, able to consistently provoke the brain wave. We propose the use of P300 detection to identify the scientific interest in a large set of images and train a computer with machine learning algorithms using the subject's responses to the stimuli as the training data set. As a first step, we here describe a number of experiments designed to relate the P300 brain wave to the cognitive processes related to placing a scientific judgment on a picture and to study the number of images per seconds that can be processed by such a system.

Key words: P300; Curiosity Cloning; ERP; image classification.

1. INTRODUCTION

Space missions are often equipped with several high-definition sensors, allowing to autonomously collect a potentially enormous amount of data. The bottleneck in retrieving these often precious data-sets is the on-board data storing capability and the communication bandwidth, which limit the amount of data that can be sent back to Earth. This issue is particularly severe for image data, which is usually quite demanding in terms of dimension (bits) and, since the best possible resolution and quality is normally required by scientists, also hardly compressible in size. Hence, despite the fact that explorative robots could take a vast amount of pictures, these will eventually have to be reduced in number. Separating the scientifically relevant pictures from the less relevant ones is the crucial task. Consequently, the robot has to evaluate in real-time the scientific content of a picture, i.e., to assign a "Scientific Richness Index" to each picture, and set its priority accordingly. The problem is shifted to the definition of such an index.

In 2007, the two NASA rovers Spirit and Opportunity received an update which made them able to detect dust-devils in the martian landscape [CFB+08]. This constituted the first onboard science analysis process on Mars, and so far the only example of selective data acquisition by exploratory rovers. The algorithm (still in use) is essentially based on the detection of changes between subsequent pictures and works well whenever the acquisition campaigns are run in still conditions. The picture interest is thus related to the "amount" of moving objects in the picture itself.

Classifier systems based on supervised learning could be used as a more general alternative. These systems could learn a possible dependence between picture features and the *subjective* scientific interest of a picture as evaluated by a given expert. This is information that is difficult to extract reliably. The expert (e.g. a geologist with an expertise in Martian rocks analysis), needs to evaluate hundreds, even thousands of pictures and to rate each one of them one by one to define their scientific interest.

Here we propose to extract this 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 related work performed by Gerson et al. [GPS06]. The main potential features of the proposed approach are the faster rate at which pictures can be presented to the scientists with respect to an interview-approach, and the reliability of the classification that could potentially be much higher. Looking for interesting features is looking for the inexplicable, highly unusual, or odd. In other terms, scientific interest is associated with the picture's features which arouse speculation, interest, or particular attention.

It is well known from neurophysiological studies that when we look at images which arouse such mental responses, our 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 (see [HM02] for a good introduction to the P300 wave). The P300 as an event-related potential (ERP) shows interesting features: its magnitude is associated with the level of attention the stimulus arouses, it cannot be fine controlled, and it is reported to be, at least partially, independent from consciousness. We aim to demonstrate that correlating the level of attention with the corresponding sensorial stimulus, it is possible to assign a scientific interest level to the stimulus presented. Moreover, since the P300 shows attention arousal at its very beginning, it is possible to classify the interest-level of an image quicker than by directly interviewing the subject, and removing any bias operated by the subject's conscious filtering. For a large set of images, as it is required to train a computer with machine learning algorithms, reducing the time dedicated to the analysis of an image can have drastic effects on the total time required of the subject to spend "looking at images".

In our vision, with the data set obtained by evaluating P300 signals associated with each picture, they are later used to train and test a classifier which ideally reacts to stimuli showing the same level of scientific attention that had been monitored from the scientists. In short, scientists' scientific attention would somehow be replicated - or "cloned" - into an artificial system.

This paper contains the description and preliminary results of experiments performed during the first half of 2009 aimed to prove the feasibility to reliably and quickly extract the scientific interest on images presented to a scientist utilizing RSVP.

2. EXPERIMENTS

Each experiment described was carried out independently by two groups located in different premises. Care was taken to replicate the experimental environment as accurately as possible and the ITU-R BT. 500-11 recommendation [BT02] was used as a baseline. To ensure equal experimental set-ups an image visualization software named Curiosity Cloning Viewer (*CCViewer)* was developed [Ruc08] and used throughout the project. The software has been released under BSD license and can be downloaded from the internet at `sourceforge.net/projects/ccviewer/` The two set-ups differed solely in the EEG recording apparatus used. The first group operated at Dublin City University (DCU) in Ireland and used two 2 channel devices with a sampling rate of 254 samples per second and a 12-bit sampling resolution. Electrodes were placed at Pz, Cz, P3 and P4 according to the international 10-20 system. A joint mastoid reference was used between both of the EEG devices which when joined provided 4 sampling channels. A ground electrode was placed on the chin of the participant. In parallel, a second group operated at the Swiss Federal Institute of Technology (EPFL) in Lausanne, Switzerland and recorded with 36 channels device. The EEG signals were acquired at 2048 Hz and 24-bit sampling rate from 32 electrodes that were placed on the scalp of the subjects according to the 10-20 international electrode positioning system. A Biosemi Active Two amplifier was used for amplification and analog to digital conversion of the recorded EEG signal. The different EEG set-ups were selected to prove both the possibility to perform such experiments with an extremely compact and portable device (4 channels) and to make sure to be able to access all possible relevant brain activity.

(a) (b)

Figure 1. Examples of an oddball (a) and non-oddball (b) images used for Phase 1.

2.1. Phase 1 - Calibration, Presentation Rate, Subconscious Perception and Learning

The aim of the first phase of the experiments was multifold. The most basic objective was to confirm that the P300 signal can be reliably detected with the used experimental set-up and with the available tools. The next goal was to analyse how P300 detection reliability is affected by the rate of the image presentation (i.e. the number of images presented per second). Then it was checked if P300 activity is evoked also in situations when the image presentation rate rules out concious perception of visual stimuli. Finally, the impact of the learning effect on the detection of P300 was assessed.

In order to fulfill these objectives, the classical oddball paradigm [HM02] has been used throughout the first phase of experiments. Visual stimuli consisted of a subset of 3204 images of grey stones luminated with a uniform ambient light. 25 of those images contained in addition to the stones, a sand model of a spacecraft, thus constituting 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 first phase experiments are given in figure 1.

The first phase was divided into 4 experiments each related to one of the aforementioned scientific goals. Every experiment involved the presentation of one or more image sequences to experiment subjects. The 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

Table 1. Parameters of the Calibration experiment

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 2. Parameters of the Presentation Rate experiment

were instructed to count the images containing the spacecraft model and were made familiar with examples of an oddball and non-oddball image. After that, the actual sequence of the images was presented with the EEG signals being recorded, always preceded by a countdown screen of duration 5 seconds that allowed the subjects to prepare for the experiment, reducing the surprise effect of the sequence start.

The parameters of the first experiment, further referred to as the *Calibration* experiment, are summarised in Table 1. The goal here was to verify that the experimental setup allows for a reliable P300 detection. The experiment involved 4 subjects, and 5 different sequences of images. Each of these sequences consisted of 40 images, 4 of which were oddball images. Oddballs were placed randomly in the image sequence. The experiment was repeated twice (and with the same 5 sequences) for each subject after an arbitrary rest period. Every image was presented to the subject for 500 milliseconds (Image Display Period, IDP), after which a neutral background appeared for another 500 milliseconds (Inter Image Period, IIP), resulting in a one image per second presentation rate. Thus, the presentation of one complete image sequence in this experiment took 40 seconds. The relatively low image presentation rate in this experiment should allow a very reliable detection of the P300 signal.

The second experiment was aimed at understanding how fast the images can be presented to the subjects while still registering a P300 response. The parameters of this experiment, further referred to as *Presentation Rate* are presented in Table 2. 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 stayed equal to 40 seconds. The number of oddball images present in the sequence was adjusted accordingly, so that the ratio of the number of oddball images to the number of non-oddball images was kept on the same level (10%). The oddballs were placed randomly in the sequences. As for the first part of the experiment all parameters are identical to the ones used in the Calibration experiment and the results of the latter were re-used.

The third issue addressed in this phase of experiments was to check that brain activity can be detected and related to oddballs even when the image presentation rate is too high to allow concious perception. Thus, a much higher image presentation rate than in the first two experiments has been used, and no inter-image blank was used (IIP=0). Two timing options have been used, resulting in displaying 30 and 60 images per second respectively, which is higher than the commonly agreed threshold of concious perception, being 20 images per second [HKW$^+$03]. For these two options, 10 different image sequences have been used, each of them containing exactly one oddball image (this fact however was not known to the subject). The oddball image placement was random, however it was enforced that it is placed within the first third of the sequence for 3 out of 10 sequences, within the middle third for 4 out of 10 sequences and within the last third for remaining 3 sequences. All parameters of this experiment further referred to as *Subconscious Perception* are summarised in table 3.

No. of subjects	*No. of sequences*	*Images in seq.*	*Oddballs in seq.*	*Repetitions*	*IDP/IIP (ms)*	*T (s)*
4	10	300	1	2	33.3/0	10
4	10	600	1	2	16.7/0	10

Table 3. Parameters of the Subconscious Pereception experiment

No. of subjects	*No. of sequences*	*Images in seq.*	*Oddballs in seq.*	*Repetitions*	*IDP/IIP (ms)*	*T (s)*
4	5	100	10	5	100/100	20

Table 4. Parameters of the Learning experiment

Finally, the issue of learning the image sequence by the subject in the case of a subsequent presentation of the same image sequence, and its impact on ERP detection was addressed. In this experiment, further referred to as *Learning*, a slightly different protocol than in previous ones was used. Each of the subjects was shown 5 different image sequences, but each one of them was repeated 5 times one time after another. Moreover the subject was made aware of this fact in advance, being also instructed that “the same image sequence is going to be repeated 5 times”. Relatively high image presentation rates have been used in order to allow the subjects to make mistakes and thus observe the learning effect, if present. All parameters of this experiment are given in Table 4.

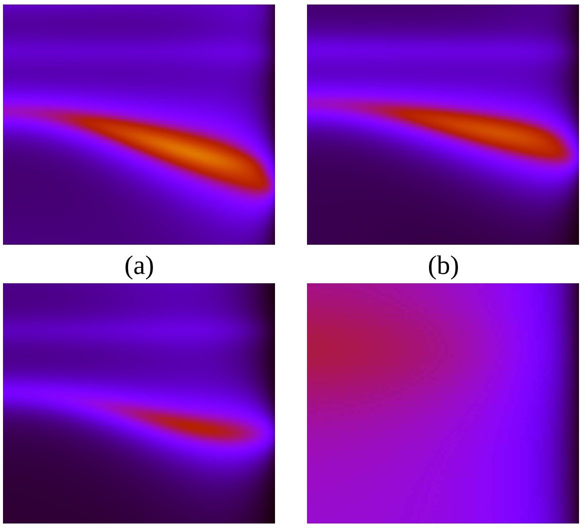
2.2. Phase 2 - Scientific Expertise

The second phase of the experiments aimed to answer questions concerning the relation between ERPs and expert knowledge or scientific curiosity. In order to meet these objectives, a special set of visual stimuli has been used, as well as two types of experimental subjects – a person who has profound scientific knowledge about the stimuli and non-experts.

The visual stimuli used in the second phase of the experiments were taken from the European Space Agency’s database of “multilayer coatings for thermal applications” [1]. The database contains images obtained dur-

---

[1] The database can be visited at the link `www.esa.int/gsp/ACT/nan/op/bigrunresults.htm`



(a) (b) (c) (d)

Figure 2. Examples of a target (a), obvious oddball (b), non-obvious oddball (c) and background (d) images used in the first experiment of the second phase.

ing the process of designing a multilayered material exhibiting predefined thermal emissivity profiles (which are called *targets*). Spectral directional properties of a material can be presented as 2-dimensional contour plots with axes representing angle and wavelength parameters and with the colour of the point representing the magnitude of the target parameter (for example emittance). Different materials, including the ideal target solution, correspond to different plots which appear as different 2-dimensional contours. However, as a material matching exactly the desired properties is not obtainable, the best found solution will only be similar to a certain degree to the ideal target solution. This "degree of similarity" is related to a simple pattern matching process (e.g. the image looks similar to the target image) in non-expert subjects, and to more complex cognitive processes in the expert (e.g. consideration on the physics of the emissivity profiles, experience of what can be considered a good match for the emissivity pattern). The image sets used were taken from different optimisation experiments for different desired ideal properties of the material and for solutions of different quality. The contours were plotted in a normalised range of parameter values and stripped from the axes and the legend.

In this phase, two experiments were conducted. The first one, called *Expertise* was designed to find out if there is a difference in P300 responses between subjects who possess scientific knowledge about presented stimuli and non-expert subjects. The experiment used a modification of the oddball paradigm, with two types of oddballs: obvious and non-obvious. In each session, the non expert subject was presented an image corresponding to the target solution and instructed to "look for similar images". The subject was also shown an example image considered an obvious oddball in order to be informed about the amount of acceptable differences between target solution and "good" solutions. Then a sequence of images was presented, which contained plots of materials with properties different from the ideal target (background images), very similar to the target (obvious oddballs) and slightly similar to the target (non-obvious oddballs). Examples of such images are shown in Figure 2.1, whilst the parameters of the experiment are presented in Table 5.

No. of subjects	*No. of targets*	*No. of sequences per target*	*Images in seq.*	*Oddballs in seq.*	*Repetitions*	*IDP/IIP (ms)*	*T (s)*
4+1	2	5	50	3+3	2	500/0	25

Table 5. Parameters of the Expertise experiment

No. of subjects	*No. of sequences*	*Images in seq.*	*Oddballs in seq.*	*Repetitions*	*IDP/IIP (ms)*	*T (s)*
1	5	50	10	2	750/0	37.5

Table 6. Parameters of the Curiosity experiment

In total 5 subjects were used, 1 expert (the European Space Agency's scientist conducting the forementioned study on multilayered materials) and 4 non-experts. Two different target images were used, with 5 image sequences prepared for each of them. Every sequence contained 3 obvious and 3 non-obvious oddballs. As in previous experiments, every measurement was conducted twice. A moderately fast image presentation rate without the Inter-Image Period was used, which resulted in sequences of 25 seconds in length.

The second experiment of phase 2 to which we will refer to as the *Curiosity* experiment, was conducted on the expert subject only. No target image has been used. Non-interesting background images were mixed with potentially interesting oddball images selected by researchers preparing the image sequences, and which represented material properties that may evoke a subject's curiosity. The subject was instructed to "look for interesting properties in the displayed images". Parameters of the experiment are shown in Table 6. Differently from the Expertise experiment, the (expert) subject is no longer asked to perform pattern matching. Instead, with this experiment we wish to assess the potentiality of a subject's scientific curiosity being imprinted on his brain wave activity. Should we be able to subsequently train an artificial system that displays similar curiosity and attention properties to the ones of the scientist, that machine would be able to look for scientifically interesting features in images in the same way the scientist would. A visionary scenario could thus include a robot on Mars evaluating images by using the scientific curiosity of certain scientists back in earth which it has learned to imitate.

3. DISCUSSION AND PRELIMINARY RESULTS

For the generation of the DCU Figure's presented here, a bandpass filter was applied to filter out frequencies outside of .1hz to 18hz. The averaged signal diagrams shown here are for site Pz since it showed the strongest differentiated signal between stimulus and non-stimulus images.

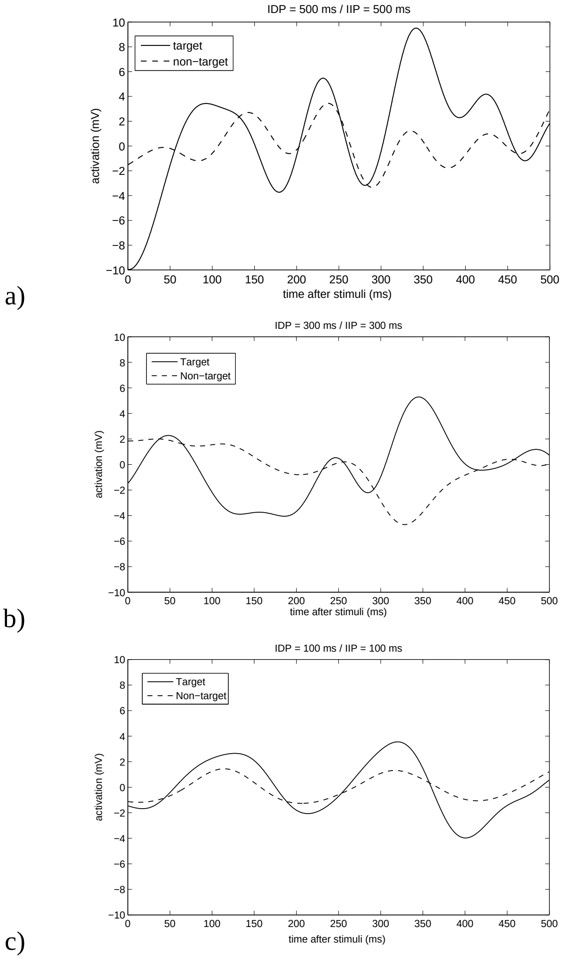


Figure 3. Averaged results for the Presentation Rate experiments of Phase 1 (EPFL)

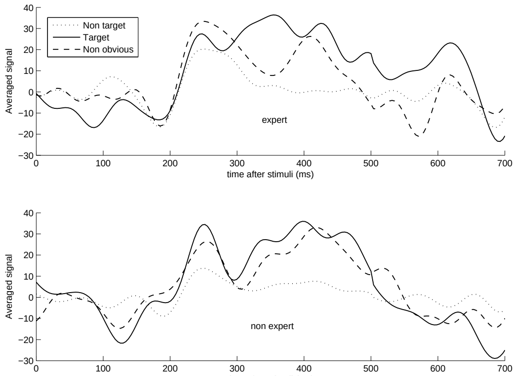


Figure 4. Averaged results for the Expertise experiment of Phase 2 (DCU)

During all the experiments, a significant amount of data was recorded. A full and complete analysis of all the acquired data will require some time and will be presented in separate works aiming at assessing the possibility of classifying the signals and using the trained classifiers to identify, in the *Curiosity* experiment, those images (i.e. materials) that the expert thought to have a promising scientific application. Here we present the preliminary analysis on the averaged data from the *Presentation Rate* and the *Expertise* experiments aimed at proving that the recordings do contain the signature of the diverse subject cognitive activity after the presentation of the different stimuli and that such an activity is different in expert and non expert subjects when scientific expertise does influence the picture judgment significantly.

Consider the average signal after the presentation of the stimuli during the *Presentation Rate* experiment. We also average the data across the different subjects and in both repetitions of the experiment. Referring to the results plotted in Figure 3, which come from the experiments carried out in Switzerland, we clearly see in all averaged curves a “positive deflection in voltage at a latency of roughly 300 ms” that is the event related potential (ERP) commonly named P300. In this particular experiment the cognitive function detected is related most probably to the decision making process the subjects undergo when deciding if the picture is an odd-ball. It is interesting to note that the magnitude of the P300 seems to decay for faster visual stimuli rates. When the image display period and the inter image period get faster we note the presence (see the averaged non target signal for IDP=100ms in figure 3c) of a visually evoked potential having the same frequency as the IDP. To this signal some information on the cognitive activity is added in the averaged target signal so that we still see a faint footprint of the P300 wave in the augmented amplitude of the oscillation at around 300 ms. Whether or not this can be used to extract conclusions on single non averaged signal remains to be determined in a more detailed analysis.

Consider now the averaged signal after the presentation of the stimuli during the *Expertise* experiment. Here we differentiate between expert and non expert subjects, and we average across the signal following a target stimuli, a non target stimuli and a non obvious stimuli. We remind the reader that in the latter case a profound scientific expertise can help in discriminating it from the target stimuli. The results of the averaging process are shown in figure 4 and come from the experiments that took place in Ireland. In these plots we clearly see the P300 wave (remember that IDP is here 500 and IIP is 0) after a target stimuli (obvious oddball) in both expert and non-expert subjects.

Interestingly, non-expert subjects seem, on average, not to discriminate between obvious and non-obvious stimuli. That is, both obvious and non-obvious matches evoke a very similar potential (P300 wave). This observation is not surprising since, after all, these subjects lack scientific expertise on the content of the presented images, and are just performing pattern matching. The expert subject, on the other hand, has a different reaction (after an obvious and a non-obvious stimuli) recorded distinctly in the EEG signal around the 300ms region, since the high peak observed for obvious matches does not appear in the case of non-obvious matches. We interpret this difference as a result of the different knowledge leveles among subjects, implying that the expert subject did not classify the non-obvious oddballs as matches to the initially presented target. Rather than performing pattern matching in order to place a judgment on the interest of the stimuli, the expert subject performed a more profound cognitive analysis of the image's scientific content which is shown in his brain signal.

Of note in the production of Figure 4, specifically graph (b), the average signals shown in Figure 4(b) are computed across 4 subjects, where each line shown is an average of the average signal for each subject for that particular image class (obvious, non-obvious, background). Whilst this data is beneficial for showing a summary of our findings, and providing an indication of what can be utilized, due to the averaging process the data presented must be interpreted with caution. Nevertheless, as an indicator for further detailed work, they do provide a strong indication of a typical subjects EEG response to a particular class of image.

4. FUTURE WORK AND CONCLUSIONS

In this article, we have introduced a new idea for collecting expert knowledge. We propose to skip the long and tiring interviews that are sometime necessary to access scientist expertise by directly accessing their brain waves.

More in detail, in our experiments the P300 brain wave is detected in a rapid serial visualization experiment with image display periods that can be as short as 100ms. The presentation speed effects the intensity of the P300 and a careful balance between speed and wave amplitude needs to be found when analysing large amounts of pictures. The averaged analysis of the signals recorded during the experiments described highlights the presence of a cognitive activity related to the picture judgment that, most important, is deeply influenced by the subject scientific expertise on the image significance. These results prove that it makes sense to train classifiers able to extract in real time scientifically interesting features in images, based on classifications performed previously by experts. Training of such classifiers and the assessment of the number of images per second such a system would be able to analyse reliably are the subjects of our future efforts.

Finally we would like to make some notes with the interpretation of the preliminary analysis shown in this paper. The figures presented are designed to demonstrate evidence of a strong indication of a subject's response to stimulus, and the potential to derive expert opinion from domain specialists. In and of themselves, these figures should not be used for conclusive proof of our ability to extract this domain knowledge from expert subjects, but rather provide an indication that this distinction does exist and could be exploited. Primarily this is because the figures we present in this paper are the average of the average signal for each subject. This is useful for presenting the differences in waveforms that can exist, but the construction of discriminative classifiers clearly requires individual training samples rather than averages for their effective construction and hence utilization of domain knowledge.

As such, the utilization of our measured signals for the purposes of machine learning requires extensive clean-up operations, including bandpass filtering, signal transformations, detection of erroneous signals (such as eye blinks) and so forth in order to provide a clean data set on which a classifier can be constructed. Furthermore, we view it as a significant challenge to build a generalized classifier which can work across subjects of similar domain knowledge. Our experiments have demonstrated that each subject can present quite unique waveforms, and that whilst construction of individual classifiers coupled to a single subject is readily obtainable, the task of a more generalized classifier which leverages multiple subject's EEG measurement is a considerable challenge. This is partly due to the complexity of the presented images and the variable latency of ERP components (recognition of stimulus may not be time locked). Individual differences also occour on a per trial and a per subject basis, involving the latency and amplitude of the detected ERP components.

5. ACKNOWLEDGMENTS

The authors wish to acknowledge the European Space Agency Ariadna scheme (`www.esa.int/gsp/ACT/ariadna/index.htm`) for having initiated and supported this research.

REFERENCES

[BT02] I.T.U.R.R. BT. Methodology for the subjective assessment of the quality of television pictures. *International Telecommuncation Union, Geneva, Switzerland*, 2002.

[CFB$^+$08] 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*, 19(5):467–482, 2008.

[GPS06] 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*, 14(2):174–179, 2006.

[HKW⁺03] M. Hoshiyama, R. Kakigi, S. Watanabe, K. Miki, and Y. Takeshima. Brain responses for the subconscious recognition of faces. *Neuroscience research*, 46(4):435–442, 2003.

[HM02] T. Hruby and P. Marsalek. Event-related potentials-the P3 wave. *Acta Neurobiologiae Experimentalis*, 63(1):55–63, 2002.

[Ruc08] M. Rucinski. Curiosity cloning image viewer user's manual. Technical Report CCIVUM01, European Space Agency, the Advanced Concepts Team, 2008. Available on line at http://www.esa.int/act.