ERP-based sustained attention type classification using Hilbert-Huang Transform: The bottom-up and top-down paradigms

A. Georgiadis, E. Kasimidou, K. Zafirouli, K. M. Bintsi, Ch. Papadaniil, and L. J. Hadjileontiadis

Aristotle University of Thessaloniki

Abstract— **The purpose of this analysis is to study the cognitive reactivity in different stimuli examining the ERPs in frequency domain using empirically based data-analysis method. More specifically, the effects of sustained visual attention performance on Event Related Potentials (ERPs) when distinguishing between bottom-up and top-down processes are investigated in this paper. The N170, P300 and N400 are the most significant ERPs for the sustained attention analysis and applying appropriate time windows on the averaged 256-channel EEG signals is required. Following a specific procedure, the Hilbert-Huang transform is applied on each waveform for the extraction of the feature matrix of each subject, which is the input in the Random Forest classifier. The results show that the selection of specific channels, that are associated with sustained attention, provide a better accuracy than using all channels. Moreover, by separating the subjects according to gender an accuracy of 73% is reached. Finally, the division between fast and slow subjects with respect to response time provides the maximum classification accuracy of 78%.**

Keywords- **Sustained attention, bottom-up, top-down, Hilbert-Huang Transform, ERPs, EEG, random forests, classification**

I. Introduction

Attention and vigilance are significant characteristics of modern people, as the way of living demands concentration on a great amount of information during daily activities. Attention is the brain function responsible for the activation of cognitive processes in order to focus on stimuli. Out of the four types of attention - sustained, selective, alternating and divided - this paper examines sustained attention, or vigilance, which is reflects the ability to concentrate on a specific task for prolonged period of time without distraction. The main processes of sustained attention performance are the top-down and bottom-up. The top-down type describes the mechanism that achieves the discrimination between the signal-target and noise using previous acquired knowledge. On the other hand, the bottom-up perspectives describe functions that depend on the characteristics of the signal and explain the subject's ability to detect the target stimulus.[1]According to previous research these two attentional systems are uncorrelated and independent.[2]

The purpose of this study is the classification of EEG signals, that are the result of bottom-up and top-down processes, so to reveal the two processes via the reflected brain activity.

A. Event Related Potentials (ERPs)

An ERP is the formulaic response to a stimulus, such as a specific sensory, cognitive, or motor event, measured using electroencephalography (EEG).[3] It is considered as a non-invasive means of evaluating the function of the brain and transcends behavioral procedures, as it provides a continuous measure of processing between a stimulus and a response, allowing us to observe the temporal analysis of the signal. ERP waveforms consist of a series of positive and negative voltage deflections, which are referred to by a letter (N/P) indicating polarity (negative/positive), followed by a number indicating the latency in milliseconds. Specifically, in our research, more information on sustained attention seems to be provided by N170, P300 and N400 waveform components. In particular:

1) The N170 is a component of the ERP that reflects the neural processing of faces. [4] Particularly, recorded by EEG, when potentials evoked by images of faces are compared to those elicited by other visual stimuli, the former show increased negativity 130-200ms after stimulus presentation, with the response being maximum over occipito-temporal electrode sites.
2) The P300 wave is an ERP component elicited in the process of decision making. It is regarded as an endogenous potential, as it links not to the physical parameter of a stimulus, but to a persons reaction to it. Using EEG, P300 appears as a positive deflection in voltage with a latency of 250 to 500ms and it is measured most strongly by the electrodes covering the parietal lobe.[5]
3) The N400 wave is part of the normal brain response to meaningful stimuli, such as signs and pictures. It is a negative-going deflection that peaks around 400 ms post-stimulus onset, although it surfaces with a latency of 250 to 500ms, and is typically maximal over centro-parietal electrode sites. [6] [7]

B. Previous Research

Over the past few years, a significant amount of research has been conducted in the field of sustained attention, in order to study the cognitive reactivity in specific stimuli (visual, auditory, somatosensory).[8][9] The recording of brain electrical activity is acquired and analyzed using the ERPs. In visual attention, which is studied in this paper, ERPs are used to investigate the difference between top-down and bottom-up brain response.[10][11] Through ERPs, the different regions of cortex that are activated according to visual stimulus are examined. Moreover, using the ERPs frequency components, the relationship between sustained attention processes and brain rhythms is studied.[12]

Traditionally, EEG data was processed by use of the Fourier Transform (FT) or the Wavelet Transform (WT).[13][14] The present work, however, experiments with the newer Hilbert-Huang Transform (HHT) for the analysis of the data since it may better reveal the spectral and temporal properties of a signal than the conventional methods.[15]

II. Materials And Methods

A. Subjects

Experimental data correspond to 16 subjects (7 males, 9 females) with mean age of 22 ± 0.5 yrs. All subjects had normal color vision and no history of neurological problems. The measurements were made using the EGI Clinical Geodesic EEG System 400, which allows simultaneous recording of 256 channels, at 250 Hz sampling frequency.

In order to ensure satisfying measurements, the electrodes used for EEG were sunk in fluid, so as to achieve good contact with the scalp of the subject. All subjects have been informed not to consume caffeine before the experiment, in order to avoid generation of additional noise and its interference with the desired signal. Moreover, they all gave their consensus prior to their participation in the experiments.

B. Stimuli and Tasks

The stimuli consisted of 125 isosceles triangles each with a different color and orientation combination.

In our experiment, subjects were firstly asked to concentrate their gaze on a cross at the center of a PC screen. After that, a randomly selected sample, a colored and oriented triangle appeared, for some milliseconds, and that was the target for the current trial. After the target presentation, there was a short delay on screen, consisting of the fixation cross; then an array showed up, consisting of the target and three distractors in the four quadrants of the screen. The target was randomly located in one of the quadrants.

The distractors were selected in such a fashion as to create either bottom-up or top-down conditions. In the bottom-up instances, the color and orientation of the three distractors differed from that of the target. In the top-down condition, however, only the orientation of the distractors differed from that of the target. The positions of the triangles were invariable, located on the upper-left, lower-left, upper-right, and lower-right, respectively. The array remained on the screen until the participant made his/her choice. This was followed by another time period of fixation on the cross, after which a new trial began.

Subjects sat in front of the computer screen and were advised not to move or talk during the trials. Each trial of the task included a target. Half of trials were of the top-down condition and half were of the bottom-up condition. Subjects were required to determine whether the target was to be found on the left or the right half plane of the screen, using their peripheral vision. Noting that before starting the experiments, participants were subjected to two explanatory test trials, in order to learn and understand the task.

C. EEG Data

Experimental data elicited from each subject were included in a registry of segmented files of size 275x375. The lines 1:256 represent the 256 channels, while line 257 the reference channel and 375 constitutes the time samples. Each segment has a time interval of 1500ms. The segmented files could belong in one of the following categories:

- Target (tar): the EEG response to the observation of the target-triangle.
- Bottom-up (bot): the EEG response to triangles of different color and orientation.
- Top-down (top): the EEG response for the case of triangles with different orientation only.

The preprocessing analysis included bandpass IIR Butterworth 6th-order filtering at 0.3-30Hz, artifact detection, bad channel replacement, baseline correction and segmentation. To extract the ERP signal from the background EEG noise, averaging was used. Two matrices of ERP waveforms were produced for each one of the 16 participants, one as the average of all bottom-up experiments and the other for all top-down experiments The whole EEG and ERPs data analysis was performed using Matlab 2015 (The Mathworks, Natrick, USA).

D. Empirical Mode Decomposition (EMD)

EMD is an adaptive method that decomposes a signal into a finite and often small number of components, called Intristic Mode Functions (IMFs), without leaving the time domain[16]. IMFs form a complete and nearly orthogonal basis for the original signal, and, therefore, are sufficient to describe it. An IMF is defined as a function that satisfies the following requirements:

- The number of extrema and the number of zero-crossings must either be equal or differ at most by one.
- The mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The fact that these functions are all in the time-domain and of the same length as the original signal allows for varying frequency in time to be preserved. Obtaining IMFs from real world signals is important, since natural processes often have multiple causes, and each of these causes may happen at specific time intervals. This is a great advantage of EMD over Fourier and Wavelet transforms.[17]

E. Hilbert Spectral Analysis

Hilbert spectral analysis is a signal analysis method applying the Hilbert transform to compute the instantaneous frequency of signals according to $\omega = \frac{d\theta(t)}{dt}$. After performing the Hilbert transform on each signal, the data can be expressed in the following form:

$$X(t) = \sum_{j=1}^{n} a_j(t) \exp\left(i \int \omega_j(t) dt\right).$$

The above equation gives both the amplitude and frequency of each component as functions of time. This frequency-time distribution of the amplitude is designated as the Hilbert spectrum $H(\omega, t)$.

F. Hilbert-Huang Transform (HHT)

HHT is an empirically based data-analysis method.It is the result of EMD and the Hilbert spectral analysis (HSA).The HHT uses the EMD method to decompose a signal into IMFs, and applies the HSA method to the IMFs to obtain instantaneous frequency data. Since the signal is decomposed in time domain and the length of the IMFs is the same as the original signal, HHT preserves the characteristics of the varying frequency.[16]

G. Feature Extraction

Each ERP was subjected to HSA. The result is the HHT of each ERP.

Six features have been chosen for the purpose of classification performed. In particular:

1) The global maximum value of instantaneous amplitude of HHT.
2) The instantaneous frequency at which the maximum value is observed.
3) The standard deviation of frequencies centered around the peak, which corresponds to values that exceed 10% of the maximum value.
4) The second maximum value of instantaneous amplitude of HTT
5) The instantaneous frequency at which the second maximum value is observed.
6) The standard deviation of frequencies centered around the peak, which corresponds to values that exceed 10% of the maximum value.

Following the above procedure for all subjects, a feature matrix was extracted for each Case Study.

H. Classification

The classification goal of our study was to accurately classify sustained attention types into bottom-up and top-down type using the six features extracted from the HHT analysis. For cross validation purposes, the hold-out method was used. Data were separated in training (75%) and testing sets (25%). As far as classifiers are concerned, the Random Forest technique with 200 trees was used.

Our study consisted of the following Case Studies:

1) *Case Study 1: All channels*
   At first, all 256 channels were utilized in a universal 100-500 ms window - that contains all three ERPs of interest - and features were extracted for each channel, for each participant and for each class.
2) *Case Study 2 and 3: Selected Channels - FP and F*
   The second case study works with a reduced number of channels since the frontal (F) and parietal (P) cortical areas are primarily involved in the mediation of sustained attention [18]. In addition, given that all subjects were under the age of 25, it is expected that P300 will demonstrate higher intensity in frontal regions of the brain [19]. The channels were, subsequently, organized into two groups; the first covering the fronto-parietal (FP) areas - Case Study 2 - and the second including only channels corresponding to the frontal lobe (F) - Case Study 3.

The above classification analysis initially took place for all participants, both males and females. Nonetheless, the gender factor was also considered; hence, the subjects were further categorized into males and females and Case Study 3 was performed for each category. Additionally, the impact of response speed was studied; thus, the subjects were categorized into fast and slow, based on the mean response time of each subject, and Case Study 3 was again performed for each category.

III. Results And Discussion

A. ERPs and HHT

Figure 1 shows an ERP waveform, drawn from the Pz electrode for the two types of experiments. As it is clear from Fig. 1, P300 amplitude is higher for bottom-up than the top-down experiment. In Fig. 2, the IMFs of the bottom-up signal shown in Fig. 1 are illustrated, whereas in Fig. 3 its HHT is depicted.

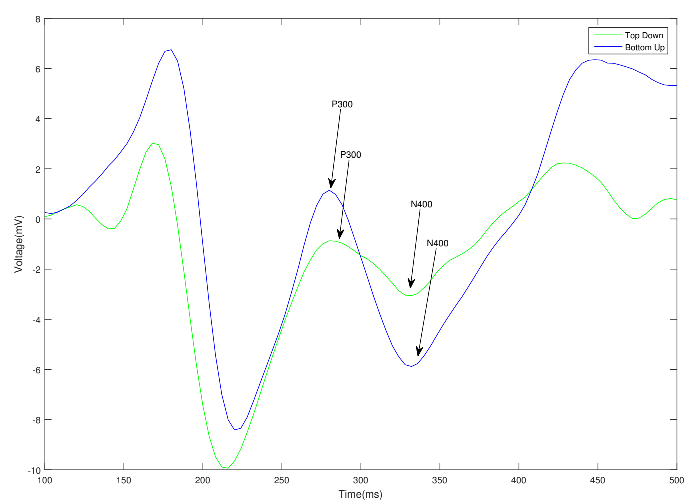

Fig. 1. ERP response at electrode Pz for bottom-up (blue line) and top-down (green line) experiments

B. All participants - Experimental Results

As expected, isolating the channels that correspond to areas of the brain associated with vigilance manages to increase accuracy up to 8%. The results are presented in Table 1.

C. Gender and Time Response Analysis - Experimental Results

For this part of the analysis only Case Study 3 is presented since it is the most efficient. (Table 2)

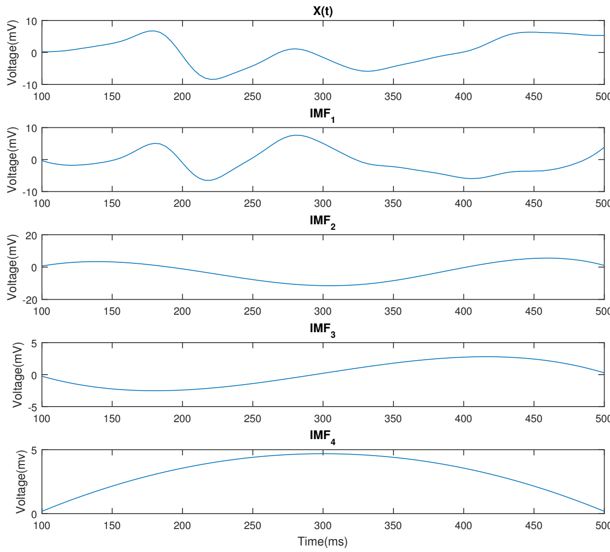

Fig. 2. Decomposition of the bottom-up signal on Pz using the EMD method

TABLE I

ALL PARTICIPANTS

	EEG Channels	Accuracy	Recall
Case Study 1	All	61.20%	0.612
Case Study 2	Fronto-parietal	62.51%	0.626
Case Study 3	Frontal	**69.38%**	**0.694**

Table 1: The classification results for all the participants. CS1 - all 256 EEG channels, CS2 - channels of the frontal and parietal cortical areas, CS3 - only channels of the frontal region of the brain (P300).

TABLE II

GENDER AND TIME RESPONSE ANALYSIS

	EEG Channels	Accuracy	Recall
All participants	Frontal	69.38%	0.694
Male	Frontal	71.43%	0.714
Female	Frontal	**72.50%**	**0.725**
Fast	Frontal	71.67%	0.717
Slow	Frontal	**78.00%**	**0.780**

Table 2: The classification results after processing independently the different categories of participants.

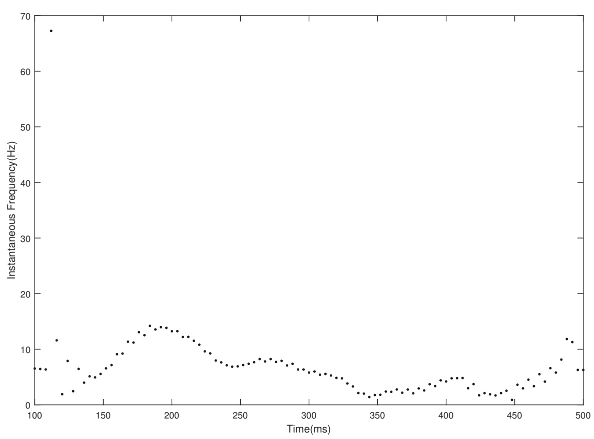

Fig. 3. Hilbert Huang Transform of the bottom-up signal on Pz electrode

By separating the analysis of the responses of the subjects according too gender, accuracy increases by 2% for male participants and 3% for female ones compared with the analysis of all participants.

Furthermore, taking into account the average time taken by each participant to give their response and carrying out a separate analysis for faster and slower participants, increases accuracy by 2% for the fast and more than 8% for the slower ones.

IV. CONCLUSION

To begin with, one can observe the improvement of classification accuracy when only specific channels are used, justifying the selection of channels that was made, focusing on those in brain regions associated with the activity of sustained attention. While the results are satisfactory, the channel selection was based on literature review assumptions on the specific electromagnetic activity of the brain, hence it is not optimal. Further tests should be carried out where different sets of electrodes, properly grouped using a selection algorithm, could be tested for potential high influence of each ERP without any *a priori* knowledge.

Moreover, it should be noted that the EMD and HHT proved to be valuable tools in joint time and frequency analysis of the ERPs, although leaving room for improvement. In their work, Huang & Wu[20] outline EMD's weaknesses: Mode-Mixing, best IMF selection and uniqueness. Moreover, a given input signal may not be decomposed into the same IMFs every time. This problem makes it hard to implement feature extraction, model training and pattern recognition since the feature is no longer fixed in one labeling index. To overcome these issues, Huang & Wu[21] proposed a new method, the Ensemble EMD. This new approach consists of sifting an ensemble of white noise-added signal and treats the mean as the final true result. Using EEMD instead of EMD is a good starting point for further research of this topic.

Overall, the findings of this study provide evidence that sustained attention is a complex cognitive mechanism that depends on contributions from different brain regions and is influenced by various factors such as gender and response speed. These issues are important for the study of sustained attention and for establishing relationships between different attention mechanisms and brain activation patterns.

AKNOWLEDGMENT

The authors would like to acknowledge the 16 subjects for their voluntary participation in the study. Moreover, the authors express their gratitude to the Cognitive Brain Signal Processing Lab at the Information Technologies Institute -

Centre for Research and technology Hellas, for providing access to the 256-channel EEG acquisition system.

Conflict Of Interest

The authors declare that they have no conflict of interest.

References

[1] Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: bottom-up versus top-down. Current Biology, 14(19), R850-R852

[2] Pinto, Y., van der Leij, A. R., Sligte, I. G., Lamme, V. A., & Scholte, H. S. (2013). Bottom-up and top-down attention are independent. Journal of vision, 13(3), 16.

[3] Luck, Steven J. (2005). An Introduction to the Event-Related Potential Technique. The MIT Press. ISBN 0-262-12277-4.

[4] Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. NeuroImage, 39, 19591979.

[5] Polich, John. "Updating P300: an integrative theory of P3a and P3b." Clinical neurophysiology 118.10 (2007): 2128-2148.

[6] C. J. Kaufman, Rocky Mountain Research Lab., Boulder, CO, private communication, May 1995.

[7] Kutas, M. & Federmeier, K.D. (In press). The N400. Annual Review of Psychology, 60.

[8] Carrasco, M. (2011). Visual attention: The past 25 years. Vision research, 51(13), 1484-1525.

[9] Coull, J. T., Frith, C. D., Frackowiak, R. S. J., & Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: a PET study of sustained attention and working memory. Neuropsychologia, 34(11), 1085-1095.

[10] Eimer, M. (1997). An event-related potential (ERP) study of transient and sustained visual attention to color and form. Biological psychology, 44(3), 143-160.

[11] Delorme, A., Rousselet, G. A., Mac, M. J. M., & Fabre-Thorpe, M. (2004). Interaction of top-down and bottom-up processing in the fast visual analysis of natural scenes. Cognitive Brain Research, 19(2), 103-113.

[12] Dockree, P. M., Kelly, S. P., Foxe, J. J., Reilly, R. B., & Robertson, I. H. (2007). Optimal sustained attention is linked to the spectral content of background EEG activity: greater ongoing tonic alpha (10 Hz) power supports successful phasic goal activation. European Journal of Neuroscience, 25(3), 900-907.

[13] Polat, K., & Gne, S. (2007). Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Applied Mathematics and Computation, 187(2), 1017-1026.

[14] Castellanos N P, Makarov V A (2006) Recovering EEG Brain Signals: Artifact Suppression with Wavelet Enhanced Independent Component Analysis, Neuroscience Methods vol.158, no. 2, pp.300-312

[15] Cong, F., Sipola, T., Xu, X., Huttunen-Scott, T., Lyytinen, H., & Ristaniemi, T. (2010, July). Concatenated trial based hilbert-huang transformation on event-related potentials. In Neural Networks (IJCNN), The 2010 International Joint Conference on (pp. 1-5). IEEE.

[16] Huang et al.(1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. R. Soc. London, Ser. A, 454, 903993.

[17] SB. L. Barnhart, The Hilbert-Huang Transform: theory,applications, development, The University of Iowa's institutional repository, 2011.

[18] Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: where top-down meets bottom-up. Brain research reviews, 35(2), 146-160.

[19] Tsolaki, A., Kosmidou, V., Hadjileontiadis, L., Kompatsiaris, I. Y., & Tsolaki, M. (2015). Brain source localization of MMN, P300 and N400: Aging and gender differences. Brain research, 1603, 32-49.

[20] Huang, N. E.; Wu Z. (2008). "A review on Hilbert-Huang transform: Method and its applications to geophysical studies,". Rev. Geophys. 46.

[21] Wu, Z., and N. E. Huang (2008), Ensemble empirical mode decomposition: A noise-assisted data analysis method, Adv. Adapt. Data Anal., in press.