Detection of Alertness States using Electroencephalogram and Cortical Auditory Evoked Potential Responses

Alaleh Rabie[1], Ahmed Al-Ani[1], Bram Van Dun[2] and Harvey Dillon[2]

Abstract— In this paper, we focus on identifying the alertness state of subjects undergoing the cortical auditory evoked potential (CAEP) hearing test. A supervised classification approach is adopted, where subjects were advised to indicate their alertness states in specified time instances. Two sets of features are considered here to represent the recorded data. The first is based on the wavelet transform of the background EEG, while the second is obtained from the peaks of the CAEP responses. The rational behind using the second feature set is to evaluate the relationship between CAEP responses and alertness levels. Obtained results suggest that the CAEP-based features are very comparable, in terms of classification accuracy, to the well-established wavelet-based features of EEG signals (79% compared to 80%). The findings of this paper will contribute towards a better understanding of CAEP responses at the different alertness states.

I. INTRODUCTION

Many researchers have studied alertness because of its impact on our ability to process information. Several alertness state detection methods have been proposed in the literature. These methods can be broadly divided into signal-based and video-based. Methods that fall in the first category use physiological signals such as the electroencephalogram (EEG), electromyogram (EMG), electrooculogram (EOG), and electrocardiogram (ECG). Some authors attempted to identify patterns characteristic of different alertness states. For example, the authors in [7] characterized reduction in vigilance by a decrease in the amplitude, quantity and frequency of the posterior alpha rhythm and increase in slow wave components. Most of the authors, however, adopted the discrete vector feature extraction and classification approach. The features were extracted from one or more of the main four physiological signals, namely EEG, EOG, EMG, and ECG [6], [5], [4], using different time-domain, frequency-domain and time-frequency domain based signal analysis techniques. These features were then fed to different classifiers, such as ANN [6] and SVM [4] to be classified in two states of alertness (alert/drowsy) [6], [4] or three states (alert/drowsy/asleep) [5]. Some of the promising video-based methods have been proposed in the literature, such as [3], [2], [1]. However, one needs to deal with a number of issues when using video-based methods, such as occlusions, target

[1]A. Rabie and A. Al-Ani are with Faculty of Engineering and IT, University of Technology, Sydney, PO Box 123, Broadway, NSW 2007, Australia Alaleh.Rabie@student.uts.edu.au, Ahmed.Al-Ani@uts.edu.au

[2]B. Van Dun and H. Dillon are with the National Acoustic Laboratories, Hearing Hub, 16 University Avenue, Macquarie Park NSW 2109 Australia Bram.VanDun@nal.gov.au, Harvey.Dillon@nal.gov.au
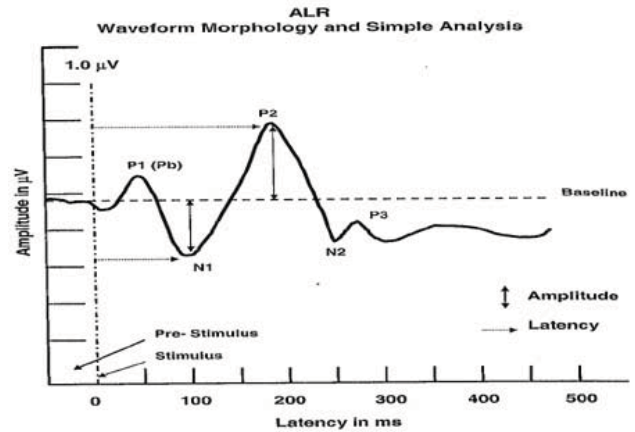
Fig. 1. Typical components of a CAEP collected from a young adult with normal hearing [8]

displacement and the large variability in eye shapes and facial expressions. In this work we will consider the first category and particularly the EEG signal, where we will process EEG data in the presence of an audio stimulus in order to obtain the cortical auditory evoked potential responses.

The cortical auditory evoked potential (CAEP) is a brain response to an auditory stimulus, originating from the auditory cortex. It has a morphology as depicted in Figure 1, and can be recorded using electrodes placed on the subject's scalp [8].

Several factors determine the morphology, i.e. amplitude and latency, of the CAEP responses, including subject age, stimulus type and the state of alertness of the subject. CAEPs are used by clinical practitioners to determine hearing thresholds and possible abnormalities of the auditory pathway. They have been the focus of renewed research [9], [10], [11]. However, the main limitation of the CAEP test is that the subject should stay vigilant or in a constant alertness state, as it has been found that the subject's alertness state during the test can significantly affect the nature of the CAEP responses [12]. Hence, when the subject shifts from one state into another, CAEP morphology changes and the averaged waveform might be annihilated. As a result it is important to keep track of the subject's state of alertness, which is not so straightforward when testing certain groups of subjects, such as children, infants, or incapacitated adults. However, the effect of the different alertness states on CAEP responses has not been fully investigated, as the literature contains a limited number of papers that deal with this issue. [13] showed that the $P2$-potential increases from

wakening subjects to drowsy as well as sleep states number two, REM-sleep and higher, while $N1$ decreases nearly to baseline level. After the change from drowsiness to sleep onset a sleep related negative potential occurs 350ms after the stimulus presentation, which is called 'sleep $N2$' or $N350$. In wakening subjects a negative potential 200ms after stimulus onset can be recorded as well but with smaller amplitude and shorter duration [14].

In this study, we will present a comparison between the performance of wavelet features that are extracted from the background EEG signal versus features obtained from the CAEP peaks (shown in Figure 1). As mentioned earlier, it has been indicated in the literature that alertness states affect the CAEP responses. Hence, the aim of this study is to investigate if it is possible to detect the alertness state from CAEP responses of normal hearing subjects. To the best of our knowledge, such study has not been conducted before.

The rest of the paper is organized as follows: in the next section we describe the data collection process. In section III, description of the two features extraction methods are presented. Classification results are presented in section IV and a conclusion is given in section V.

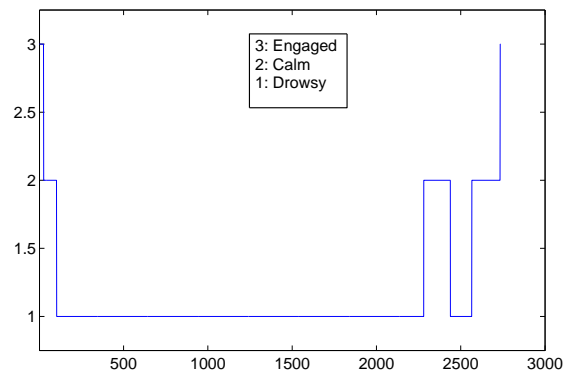## II. DATA COLLECTION AND PRE-PROCESSING

A. Participants

EEG signals were recorded from 10 adult subjects (4 females and 6 males) with normal hearing and ages that range between 24 and 53 years. All subjects had hearing levels equal to 20 dB HL or lower for four audiometric frequencies levels: 500, 1000, 2000 and 4000 Hz. Subjects were asked to try to go to sleep in a darken room during the test and press the response buttons to indicate their level of alertness each half a minute. Three buttons that indicate "engaged", "calm but alert" and "drowsy" were provided. It is important to mention that we have found a noticeable variation between the recorded labels of the different subjects, as shown in Figure 2.
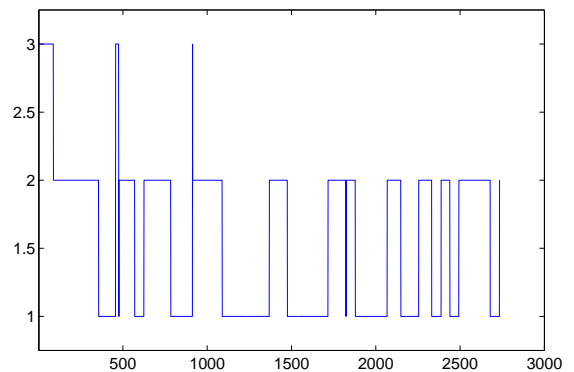
If the subject did not press any button in any of the 10-minute blocks of the 1 hour recording, then a label of 0 (sleeping) is inserted for that block. The percentage of the four classes ("engaged", "calm", "drowsy" and "sleeping") from all 10 subjects are: 15.19%, 34.29%, 46.20% and 4.32% respectively.

B. Recording Scheme and Stimulus

Data recording took place at the National Acoustic Laboratories (NAL), using a Compumedics Neuroscan Synamp 2. A multi-channel EEG Quikcap with 66 electrodes that included 64 EEG and 2 EOG channels were used to collect the data from different parts of the scalp. Each test took one hour and each stage was broken into six blocks of ten minutes with 1024 Hz sampling rate. Stimuli were presented using with an inter-stimulus interval (ISI) of 1125 ms using a loudspeaker at 55 dB SPL. The stimulus was a 21 millisecond speech sound /g/, as shown in Figure 3. The /g/ sound's primary energy has been measured in the range between 800 to 1600 Hz, and has been used in other studies of NAL [11].

(a)

(b)

Fig. 2. Recorded labels of two subjects with different fluctuation levels. Horizontal axis: time in seconds. Vertical axis: three states of alertness as indicated in the legend

C. Used Channels

Although we have collected the data from 64 EEG channels that all have been processed in another study [15], we decided to use 6 channels only in this study, namely; T7, M1 and P7 and the corresponding channels on the opposite side; T8, M2 and P8 (Figure 4). We considered those channels as they produced the maximum peaks with respect to the reference electrode that was placed on the central lobe close to Cz. A smaller number of channels can be considered; however we wanted to reduce the effect of noisy channels, as it is not unusual to receive noisy data from some of the channels.

III. FEATURE EXTRACTION

A. Wavelet-Based EEG features

The background EEG signal of each of the six channels was divided into windows of 5 seconds with an overlap of 3 seconds. The time-scale representation of each window was used to obtain 9 features that represent the energy of the dyadic wavelet transform, where an order 2 Daubechies mother wavelet was used. Thus, based on the sampling frequency, the 9 frequency bands are: 0-2, 2-4, 4-8, 8-16, 16-32, 32-64, 64-128, 128-256, 256-512 Hz. In order to form the
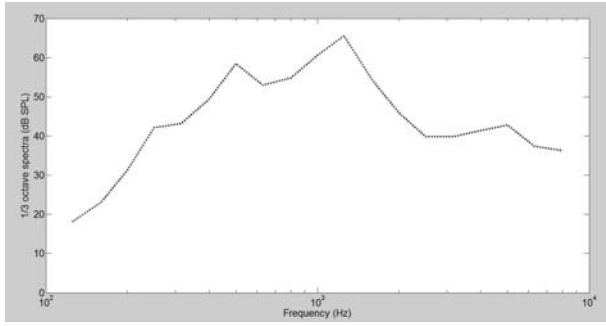
Fig. 3. Spectrum of the 21 ms speech stimulus /g/

Fig. 4. EEG cap channels location (different colours indicate different impedances on the corresponding electrodes with blue good impedance and pink bad impedance

wavelet feature vector, the six channels were concatenated which made 54 features for each window.

B. CAEP-based Features

The stimulus onset was used as a reference, were an epoch was formed using the EEG data of a given channel 200 ms before the stimulus onset and 600 ms after it. A band-pass filter was applied with cut-off frequencies of 1 and 20 Hz. To reduce the effect of baseline shift, the mean of each epoch was subtracted from its samples. 60 epochs where averaged to suppress the background EEG and emphasize the CAEP responses, as the morphology of CAEP responses may not be clearly identifiable when considering a single epoch. When a number of epochs that are aligned with reference to the stimulus onset are averaged, the background EEG is suppressed, while CAEP responses are emphasized.

An algorithm was developed to identify the $P1$, $N1$, $P2$, $N2$ and $P3$ peaks, where the search for each peak was restricted in the following time ranges from the stimulus onset: $P1 < 90$ ms, 50 ms $< N1 < 160$ ms, 150 ms $< P2 < 250$ ms, 180 ms $< N2 < 300$ ms, and 230 ms $< P3 < 400$ ms. Each peak was represented using two values; namely amplitude and latency. When no peak is found in a given region, then the latency was identified based on the minimum of the first derivative of the curve in that region. In order to obtain the same feature vector size as that of the wavelet, latencies of the five peaks and amplitude values of the first

	Engaged (T)	Calm (T)	Drowsy (T)	Sleeping (T)
Engaged (P)	0.63	0.0926	0.0322	0
Clam (P)	0.301	0.724	0.1447	0.0168
Drowsy (P)	0.069	0.1767	0.8108	0.0361
Sleeping (P)	0	0.0067	0.0123	0.9471

	Engaged (T)	Calm (T)	Drowsy (T)	Sleeping (T)
Engaged (P)	0.5696	0.0743	0.0213	0.0406
Clam (P)	0.2834	0.8001	0.1817	0.0297
Drowsy (P)	0.147	0.1256	0.7919	0.2445
Sleeping (P)	0	0	0.0051	0.6852

four ($P1$, $N1$, $P2$ and $N2$) were used. Accordingly, for the six channels, each epoch was represented using 54 features.

IV. ANALYSIS OF THE RESULTS

The obtained windows/epochs of each method were divided into segments of 30 seconds each, such that each of these segments is either used for training or testing. In other words, each segment consists of multiple windows/epochs. 75% of the segments of each subject were used for training while the remaining 25% for testing. A multi class support vector machine (SVM) was employed to individualy classify and estimate the accuracy of the labeled data for each of the ten subjects.

Confusion matrices for both methods achieved by averaging the results of the ten subjects are shown in Tables I and II. It can be noticed from the two tables that the wavelet features perform better with the engaged and sleeping classes, while the CAEP features achieved better results for the calm class, likely because CAEPs are most clearly visible in this class. Both methods achieved comparable results for the drowsy class. One important aspect that can be obtained for the two tables is that both methods tend to achieve lower misclassification rates with the increase of distance from the true class. For example, when the true class is engaged (column 2 of Tables I and II), none of the two methods predicted sleeping, and a slightly higher misclassification with the drowsy class, while the highest misclassification was achieved with calm, which is the closest class to engaged.

The overall classification accuracy shown in Table III indicates that both methods achieved comparable performance with the wavelet features achieving a slightly better performance than their CAEP counterparts. It is important

	Accuracy	STD
Wavelet Features	79.78%	8.96
CAEP features	78.81%	8.43

to mention that the developed algorithm may have not been optimal in terms of correctly identifying every peak. Incorrect values of amplitude and latency could have been corrected by considering the temporal (before and after) as well as spatial (other channels) information. Estimation of the peaks could have been also improved by averaging more epochs. However, this will increase the overlap between training and testing segments, and hence will have an influence on the generalization capability of the method. As mentioned earlier, labels were entered by subjects, where consistency was not guaranteed. This could have an effect on the obtained results. The recorded labels could be refined by a trained technician that would monitor the video of recorded sessions. This will help in reducing differences between labels obtained from different subjects and will enable us to conduct a subject-independent study.

Despite these points, the obtained results represent a proof-of-concept on the applicability of CAEP feature on the estimation of alertness levels. Hence, in addition to taking the above aspects into consideration, we are all planning to conduct a more in-depth analysis of the spatial domain by identifying the best region(s) of the brain that is(are) related to alertness identification. We will attempt in our future work to identify a set of "If then" rules that relate CAEP peaks to alertness states. Finding such rules, will be very beneficial in understanding relationships between CAEP responses and the different alertness states, and hence will add value to the utilization of CAEP in objective hearing assessment.

V. CONCLUSION

We presented in this paper an investigation on the usability of features obtained from the CAEP responses in the detection of alertness state of normal hearing subjects. EEG data was collected from 10 subjects in the presence of an auditory stimulus. Data was then processed to obtain both CAEP features and wavelet features. The obtained classification results indicate that the CAEP features have the potential to achieving comparable results to that of the well-known wavelet transform (79% and 80% respectively). Further improvement on the peak detection of CAEP responses as well as considering a different channel combination will be investigated in the future. A study on identifying relationships between the different CAEP peaks and alertness states will also be carried out.

VI. ACKNOWLEDGMENT

REFERENCES

[1] A. Varri, K. Hirvonen, J. Hasan, P. Loula, and V. Haikkinen, A computerized analysis system for vigilance studies. *Computer Methods and Programs in Biomedicine*, 39, pp 113-124, 1992.

[2] M. Patel, S.K. Lal, D. Kavanagh and P. Rossiter, Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert Systems with Applications*, 38, pp 7235-7242, 2011.

[3] M.K. Kiymik, M. Akin and A. Subasi, Automatic recognition of alertness level by using wavelet transform and artificial neural network. *Journal of Neuroscience Methods*, 139, pp 231240, 2004.

[4] M.V.M. Yeo, X. Li, K. Shen, and E.P.V. Wilder-Smith, Can SVM be used for automatic EEG detection of drowsiness during car driving? *Safety Science*, 47, pp 115124, 2009.

[5] E. Vural, M. etin, A. Eril, G. Littlewort, M. Bartlett and J. Movellan, Drowsy driver detection through facial movement analysis. *Proceedings of the 2007 IEEE International Workshop on Human-computer Interaction*, Rio de Janeiro, Brazil, 2007.

[6] Q. Ji, Z. Zhu and P. Lan, Real-time nonintrusive monitoring and prediction of driver fatigue. *IEEE Transactions on Vehicular Technology*, 53(4), pp 1052-1068, 2004.

[7] K.-A. Hwang and C.-H. Yang, Attentiveness assessment in learning based on fuzzy logic analysis. *Expert Systems with Applications*, 36, pp 62616265, 2009.

[8] J.W. Hall, *New Handbook of Auditory Evoked Responses*. Allyn and Bacon: Boston, 2007.

[9] M.P. Gorga, K.A. Beauchaine, J.K. Reiland, Comparison of onset and steady-state responses of hearing aids: implications for use of the auditory brainstem response in the selection of hearing aids. *Speech and Hearing Research*, 30, pp 130-136, 1987.

[10] H. Dillon, So, baby, how does it sound? Cortical assessment of infants with hearing aids. *The Hearing Journal*, 58(10), pp 12-17, 2005.

[11] I.M. Colrain, K.B. Campbell, Event-related potential measures of the inhibition of information processing: II. The sleep onset period. *International Journal of Psychophysiology*, 46, pp 197-214, 2002.

[12] J. Gora, I. M. Colrain, and J. Trinder. Respiratory-related evoked potentials during the transition from alpha to theta EEG activity in stage 1 nREM sleep. *Journal of Sleep Research*, 8(2), pp 123-134, 1999.

[13] T.W. Picton, S.A. Hillyard. Human auditory evoked potentials. ii: Effects of attention. *Electroencephalography and Clinical Neurophysiology* , 36, pp 191-200, 1974.

[14] C. Wen-Hung, C. Yu-Chieh, L. Sheng-Fu, Adaptive feature extractions in an EEG-based alertness estimation system, *IEEE International Conference on Systems, Man and Cybernetics*, pp 2096-2101, 2005.

[15] A. Al-Ani, B. Van Dun, H. Dillon, A, Rabie, Analysis of Alertness Status of Subjects Undergoing The Cortical Auditory Evoked Potential Hearing Test, *International Conference on Neural Information Processing, ICONIP 2012* p. 92-99, 2012.