33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Spectral EEG Featuresfor Evaluating Cognitive Load

Pega Zarjam, Julien Epps, and Fang Chen

Abstract—**This study was undertaken to investigate spectral features derived from EEG signals for measuring cognitive load. Measurements of this kind have important commercial and clinical applications for optimizing the performance of users working under high mental load conditions, or as cognitive tests. Based on EEG recordings for a reading task in which three different levels of cognitive load were induced, it is shown that a set of spectral features - the spectral entropy, weighted mean frequency and its bandwidth, and spectral edge frequency - are all able to discriminate the three load levels effectively. An interesting result is that spectral entropy, which reflects the distribution of spectral energy rather than its magnitude, provides very good discrimination between cognitive load levels. We also report those EEG channels for which statistical significance between load levels was achieved. The effect of frequency bands on the spectral features is also investigated here. The results indicate that the choice of optimal frequency band can be dependent on the spectral feature extracted.**

I. Introduction

COGNITIVE load is the amount of task demand applied on working memory when performing a mental task [1]. It is quite well-known that working memory is limited in capacity and time when holding or processing information [1]. Therefore, when a task becomes more difficult, the accessibility of working memory reduces and cognitive load increases. This may lead to decreased performance or even failed task completion, which is undesirable in any circumstance, but particularly in critical decision-making fields, such as air traffic control, fire command and military operations or when designing adaptive human computer interfaces. Thus, measuring the cognitive load experienced is a critical need [1].

Different techniques are available for measuring cognitive load; among them electroencephalography (EEG)-based methods are the most suitable for continuous and on-line assessment of cognitive load at all levels [2]. This is due to the high sensitivity of EEG to variations between different cognitive states, and task difficulty in one hand, and the easy usage of the EEG device and being less costly and bulky on the other hand [2]. Specially, the usage of EEG has become more feasible for real-world applications recently with the availability of wireless EEG systems [3].

Finding features that are good discriminators of different workloads is an important key to successfully measure and classify the cognitive load levels. Previously, a range of spectral features have been deployed for this purpose using EEG signals, indicating that the spectral features discriminate cognitive load best. This includes the use of power spectral density (PSD), the power and maximum\log power spectra [4]-[8]. But to date, other spectral features that could provide more information on the varying cognitive load characteristics have not been investigated.

This study is a continuation of the authors' previous study in which other features were employed to characterize varying behavior of cognitive load, and which identified a few frontal EEG channels as the most effective channels in separation of the three induced loads [9].

In this study, we aim to examine the usefulness of some spectral features (i.e. spectral entropy, weighted mean frequency and its bandwidth, and spectral edge frequency), and to determine in more detail the frequency band in which the maximum discrimination among the three load levels is yielded. These spectral features have been employed in past EEG medical/ brain computer interface (BCI) analysis [10]-[12], but not in the cognitive load context. For the various spectral features, the related EEG channels for which reliable information may be extracted for an EEG-based cognitive load classification system is also investigated.

II. Material

A. Experiment

EEG signals were acquired from five healthy male participants; postgraduate students aged between 24-30 years. In the experiment, the participants were asked to read text silently, which was displayed and controlled on a laptop PC with a viewing distance of 70 cm to them. The reading task was chosen to be semantically neutral and comprehension-independent to avoid any expertise effect, and assumed the reading ability of all participants to be relatively similar. To rate the task readability/complexity and ensure it induced three different difficulty loads, the Lexile framework for reading [13] was used. The test determined the different difficulty loads as 1020 for low load, 1090 for medium load, and 1150 for high load.

The task was split into three levels; participants were asked to read the displayed pages silently and pick up three (low), three and four (medium) or three, four and five (high) letter words by pressing the mouse left/middle/right button. In the baseline condition, conducted after the experiment, the participants were asked to sit relaxed and keep their eyes open. To minimize any muscle movement artifact (EMG),

Manuscript received April15, 2011. This work was supported by the Asian Office of Aerospace Research & Development, Grant No. FA2386-10-1-4029.

P. Zarjam is with the School of EE&T, University of New South Wales, Sydney, NSW 2052, Australia. She is also withATP Research Laboratory, National ICT Australia (phone: +61 2 9385 4803; fax: +61 2 9385 5993; email: p.zarjam@student.unsw.edu.au).

J. Epps is with the School of EE&T, University of New South Wales, Sydney, NSW 2052, Australia. He is also with ATP Research Laboratory, National ICT Australia(email: j.epps@unsw.edu.au).

F. Chen is with ATP Research Laboratory, National ICT Australia(email: fang.chen@nicta.com.au).

the participants were asked to sit still and their hand was placed fixed in a certain position, where they could still make finger movements for clicking the mouse buttons in response to the word stimuli. The participants were also required to refrain from blinking as much as possible during the recording to avoid ocular artifacts (EOG).

B. EEG Recording

The data used for this study consists of multi-channel EEG recordings obtained from the five consenting participants. The data were acquired using an Active Two system [14], at the ATP Laboratory of National ICT Australia in Sydney. The experiment was conducted under controlled conditions in an electrically isolated lab, with a minimum distance of 5 meters from power sources to the experiment desk, and under natural illumination. Each recording contained 32 EEG channels, according to the international 10 - 20 system. The data were recorded in a digital form, at a $f_s = 256$ Hz sampling rate. This dataset was also used in the authors' previous study [9].

III. Method

A. EEG Preprocessing and Segmentation

A notch filter of 50 Hz was initially applied to the raw EEG signals to remove the electrical mains contamination. The signals were then passed through a band-pass filter with a pass-band of 0.1-100 Hz. Visual inspection of the recorded signals showed very low ocular artifact occurrences. Therefore, no artifact removal was conducted here. The acquired EEG signals were first segmented using 5s non-overlapping rectangular windows. Each segment is denoted herein as $x[n]$ and contains $N = 1280$ samples.

B. Spectral Feature Extraction

Five spectral features were extracted from each segment of the EEG signals. They are presented as follows:

Spectral Entropy (SpEn): Spectral entropy is a measure of the distribution of normalized spectral energies, in this case within a frequency band. It is given by:

$$\mathrm{SpEn} = -\frac{1}{\log N_f}\sum_f P_f(x)\log_e P_f(x) \tag{1}$$

where $P_f(x)$ is an estimate of the probability density function (PDF) of the EEG segment amplitude spectrum. The PDF is calculated by normalizing the PSD estimate with respect to the total spectral power in each sub-band. N_f is the number of frequency bins in the PSD estimate.

According to the above equation, the spectral entropy attains its peak when all the frequency bins contain the same power. One reason for investigating this feature is that it is essentially independent of the sub-band energy, which has been shown to perform well in previous work [9].

For illustration purposes, the PSDs of the EEG signals acquired from one participant, while the three cognitive

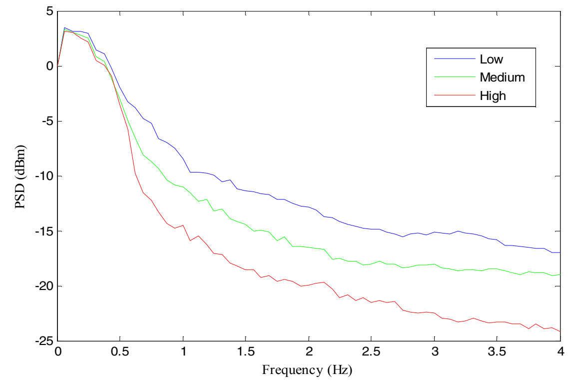

Fig.1. Extracted PSDs from the segmented EEG for participant 1 channel F8, recorded during induced high, medium, and low load levels.

loads induced are shown Fig. 1. As the cognitive load level increases, the signal power is concentrated in a smaller frequency band. Therefore, the spectral entropy decreases. The extracted entropy values in the delta sub-band shown in Fig. 1 for low, medium, and high loads are $0.6171*10^{-7}$, $0.5795*10^{-7}$, and $0.5343*10^{-7}$, respectively. This feature has been used for neonatal seizure detection previously [10].

Sub-band Energy (Enrg): The second spectral feature used here is the energy of the EEG segment in the delta sub-band (0-4 Hz) [9]. The choice of the sub-band was based on the fact that most of the energy of $x[n]$ resides below 4 Hz, which is also seen in Fig.1. In fact, energy is the integral of the signal spectral amplitudes, which is proportional to the signal power. It is an effective feature for EEG signal spatial classification in BCI applications [12].

Intensity Weighted Mean Frequency (IwMf): This feature measures a weighted mean of the frequencies present in the PSD estimate for each EEG segment [10]:

$$\mathrm{IwMf} = \frac{\sum_{i=1}^{N_{f/2}-1} p_i i df}{\sum_{i=1}^{N_{f/2}-1} p_i} \tag{2}$$

where p_i is the estimated spectral power in frequency bin of i, f_s the sampling frequency, N_f the total number of frequency bins, and. $df = f_s / N_f$.

Intensity Weighted Bandwidth (IwBw): The associated bandwidth of the IwMf feature can be calculated by [10]:

$$\mathrm{IwBw} = \sqrt{\frac{\sum_{i=1}^{N_{f/2}-1} p_i(\mathrm{IwMf} - idf)^2}{\sum_{i=1}^{N_{f/2}-1} p_i}} \tag{3}$$

Spectral Edge Frequency (EdFr): This feature is defined as the frequency below which 90% of the signal power resides. This measure has been used previously for quantifying a pathological state (i.e. the depth of anesthesia in adults or white matter injury in neonates) [11].

C. Analysis

Since the features of interest are spectral, we initially investigated the spectral components of the recorded EEG signals. It appears that 90% of the energy of the spectral components resides in the 0-3.8 Hz region, computed by the extracted EdFr. This is also confirmed visually by the PSD shown Fig. 1. Clearly, the delta sub-band practically provides the most separation between the three load levels induced in the experiment. Therefore, the performance of all features namely; SpEn, Enrg, IwMf, IwBw, and EdFr were examined in the delta sub-band and compared for all the 32 EEG channels for all participants. We initially calculated the median of all the features for each EEG channel recorded, and then compared the effectiveness of each feature using a Kruskal-Wallis test.

IV. Results and Discussion

A. Feature Comparison

TABLE I lists the EEG channels in which the features' medians calculated across all participants have shown consistent trends. It displays that the SpEn and Enrg exhibit a consistent decreasing trend as load level increases, in selected channels. However, the IwMf, IwBw, and EdFr exhibit an increasing trend as load level increases.

In order to examine how effective the extracted features are in the different load levels separation from the EEG channels, we used the Kruskal-Wallis test. The benefits of this test are: it examines more than 2 groups, is a non-parametric method, and is not affected by variations in a small portion of the data [15].The largest calculated p-values (indicating the worst case scenario) across the selected channels for each feature for all participants are displayed in TABLE II. As the p-values suggest, the Enrg feature shows a great statistical significance in differentiating the cognitive load levels in all the EEG channels. It is followed by the SpEn feature, with the second lowest set of p-values. As seen, most of the selected channels in TABLE I are confirmed in TABLE II, statistically.

TABLE I
Variations of the extracted features' medians from the EEG channels in the delta sub-band, indicating the channels that follow a consistent trend associated with the 3 load levels induced, across all participants.

Feature	EEG channels	Trend with increasing load
SpEn	Fp1, AF3, F7, CP5,P7,Pz, P8, CP6,CP2, T8, F4, F8, AF4	decreasing
Enrg	Fp1, AF3, F7, T7, CP3, P7, Pz, P8, CP6, CP2, T8, F4, F8, AF4	decreasing
IwMf	F7, FC5, T7, C3, P7, P3, Pz, P8, CP2, FC6, FC2, Fp2, Fz	increasing
IwBw	Fp1, AF3, F7, T7, C3, CP5, P3, Pz, PO3, P4, CP2, FC6, FC2, F4, Fp2, Fz	increasing
EdFr	Fp1, AF3, F7, P3, Pz, CP2, Fp2, Fz	increasing

TABLE II.
The EEG channels with a p-values<0.01 for allparticipants in the delta sub-band. The biggest calculated p-values display the worst case scenario.

Feature	EEG channels	Maximum p-value
SpEn	Fp1, AF3, T7, CP5,O1, P8, CP2,F8, Fp2	0.00357
Enrg	All 32 channels	5.0005372e-05
IwMf	P3, Pz, Oz, CP2, Fp2, Fz	0.0076619
IwBw	P3, Pz, CP2, F4, AF4, Fz	0.0091394
EdFr	P3, Pz, CP2, F4, AF4, Fp2, Fz	0.00847211

For illustration purposes, this significant statistical difference among the three load levels for the SpEn is also displayed in Fig. 2 for channel F4. As observed, the values of the SpEn feature are different for three load levels. Also, there is no overlap between the three box plots indicating complete separation among the three load levels.

B. Sub-divisions of the Delta Band

Following our investigation of the delta frequency sub-band; we divided this sub-band into finer frequency sub-bands. Therefore, the delta sub-band (0-4Hz) was split into three sub-bands of δ_0 (0-1 Hz), δ_1 (1-2 Hz), and δ_2 (2-4 Hz) using wavelet decomposition. The feature medians were recomputed for the δ_0, δ_1, and δ_2 sub-bands. It was observed that for the SpEn, and Enrg δ_0, and δ_1 sub-bands provided the same results as TABLE I, showing that the lower sub-bands seem to underlie most of the load level variations. For the IwMf, IwBw, and EdFr the medians in the finer sub-bands did not provide any significant results, indicating the load level information extracted by these features is distributed all over the delta sub-band. This suggests that the performances of the extracted spectral features are highly dependent on the frequencysub-divisions.

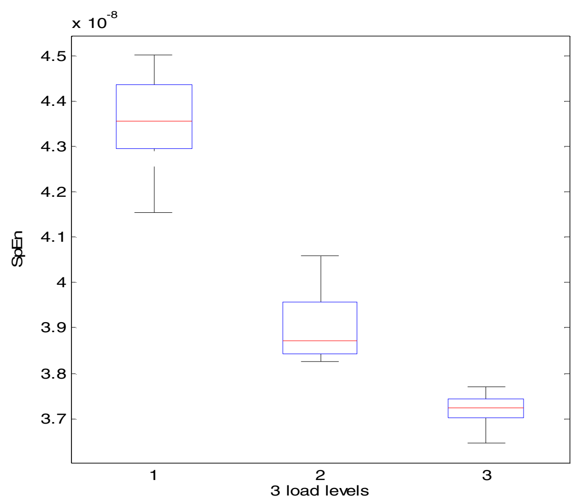

Fig. 2 Boxplot of the SpEn feature extracted from the segmented EEG data across all participants for channel F4. On each box, the red mark is the median; the edges of the box are the 25th and the 75th percentiles. Low level is denoted by 1, medium by 2, and high by 3 here.

C. Classification

To measure the classification accuracy of the extracted features, we used a multi-class support vector machine (SVM) as a classifier. The spectral features were used with three SVMs in a pair wise strategy. One SVM was used to separate low from medium, one used to separate low from high, and the third one was deployed to separate medium from high. The SVM used a linear kernel, and compared the results of the three load level classification for all 32 channels, applied on a per-subject basis. 80% of the data (for each task level for each participant) were used for training and the remaining 20% for testing. We first examined the performance of each feature individually, but it yielded low classification accuracy. Therefore, we combined all the features into the classifier which resulted in a superior performance (higher accuracy rate). This can suggest that each feature captures different information/characteristics of the signal. In other words, they complement one another in defining the cognitive load variations. The classification was performed only on 5channels with the smallest *p*-values from TABLE II. The classification results for all features, on the selected channels which are averaged across all participants are displayed in TABLE III. As shown, 3 channels, namely; Pz, Cp2, Fp2 channels present a high classification accuracy of over 95%.

V. Conclusions

In this study, we investigated the high ability of few spectral features in discriminating three cognitive load levels. This discrimination was found to be statistically significant using the Enrg, and SpEn as the best performing features, followed by IwMf, EdFr, and IwBw, respectively. The Enrg, and SpEn features trend appeared to decrease as the cognitive load level increased. Although, this trend seemed to increase for the IwMf, EdFr, and IwBw features as the cognitive load level increased. The optimal frequency band was also found to be highly dependent on the spectral features in use. The delta sub-band yielded the best performance for the IwMf, IwBw, and EdFr features. However, this highest performance was achieved for the SpEn and Enrg in the δ_0 , and δ_1 sub-bands. Clearly, this needs to be further validated on other databases. Combination of all the features into the classifier resulted in superior performance compared to one feature taken alone, suggesting that each feature captures different information/characteristics of the signal. Therefore, the feature combination defines the cognitive load variations better.

TABLE III
Accuracy of the 3 load level classification for all features by a SVM with linear kernel, averaged over all participants.

EEG channel	P3	Pz	CP2	Fp2	Fz
Classification accuracy %	86.62	95.55	95.55	97.77	82.21

We also determined the few channels which present the highest classification accuracy among the three load level induced. This suggests that a smaller number of EEG channels may be needed for future similar work. However, this needs to be further validated on a larger database. Future work includes collection of EEG data across a larger number of cognitive load levels, which will pose a substantially more difficult classification task.

Acknowledgment

The authors would like to acknowledge the volunteers for participating in the experiment, and the assistance of Mr. M. A. Khawaja for testing the designed reading task against the Lexile analyser.

References

[1] F. Paas, J. E. Tuovinen et al., “Cognitive load measurement as means to advance cognitive load theory”. In Educational Psychologist, vol. 38, no.1, pp. 63-71, March 2003.

[2] P. Antonenko, F. Paas et al., “Using Electroencephalography to measure cognitive load”. In Educational Psychology Review, DOI 10.1007/s10648-010-9130-y, Springer Science and Business MediaLLC 2010.

[3] C. Berka, D. J. Levendowski et al., “Real-Time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset”. In Intr. Journal of HumanComputer Interaction, vol. 17, no. 2, pp. 151–170, 2004.

[4] K. M. SpencerandJ. Polich, “Poststimulus EEG spectral analysis and P300: Attention, task, and probability”. In Intr. Journal of Psycho physiological Research, vol. 36, no. 2, pp. 220-232, 1999.

[5] J. Lamberts, V.D. Broek, et al., “Correlation dimension of the human EEG corresponds with cognitive load.” In Neuropsychology, vol. 41, pp. 149-153, 2000.

[6] P.F. Diez, E. Laciaret et al., “A comparative study of the performance of different spectral estimation methods for classification of mental tasks”. In the proceedings of the 30th Int' IEEE EMBS Conference, pp. 1155-1158, 2008.

[7] D. Erdogmus, A. Adami et al.,“Cognitive state estimation based on EEG for augmented cognition.” In the 2nd Int. IEEE EMBS Conference on Neural Engineering, pp. 4-7, 2005.

[8] W. Klimesch, “EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis”. In Brain Research Reviews, vol. 29, pp. 169–195, 1999.

[9] P. Zarjam, J. Epps et al., “Characterizing working memory load using EEG delta activity.” To appear in the proceedings of the 19th European Signal Processing Conference, 29 August- 2 September 2011, Barcelona, Spain.

[10] B.R. Greenea, S. Faula et al. “A comparison of quantitative EEG features for neonatal seizure detection.” In Clinical Neurophysiology, vol 119, pp. 1248–1261, 2008.

[11] O.M. Doyle, B.R. Greenea et al., “The effect of frequency band on quantitative EEG measures with Hypoxic-ischaemic encephalopathy.” In the proceedings the 29th Int' IEEE EMBS Conference, pp.717-721, 2007.

[12] Sh. Sun, “The extreme energy ratio criterion for EEG feature extraction.” In the proceedings of ICANN 2008, part II, pp.919-928, 2008.

[13] The Lexile Framework For Reading (www.lexile.com)

[14] http://www.biosemi.com

[15] M. Hollander and D. A. Wolfe, “Data statistical analysis: Nonparametric Statistical Methods.” John Wiley & Sons, Inc.