EEG-based Mental Workload Recognition Related to Multitasking

Wei Lun Lim, Olga Sourina, Yisi Liu
Fraunhofer IDM @ NTU
Nanyang Technological University
Singapore
{wlim031, EOSourina, LIUYS}@e.ntu.edu.sg

Lipo Wang
School of Electrical and Electronic Engineering
Nanyang Technological University
Singapore
ELPWang@ntu.edu.sg

Abstract— **Mental workload can be recognized from Electroencephalogram (EEG) and can be used to assess mental efforts of the user performing different tasks. In this work, we designed and implemented an experiment for mental workload recognition related to no-task, visual task, auditory task and multitask performance. The Simultaneous Capacity SIMKAP test was used to induce different levels of mental workload related to multitasking in 12 subjects. EEG data was collected with Emotiv device, processed and analyzed using power, statistical, fractal dimension (FD) features with Support Vector Machine (SVM) and k-Nearest Neighbors (k-NN) classifiers. The best accuracy of 90.39% for 2 classes and 80.09% for 4 classes using SVM was achieved when statistical and FD feature combination was used. The proposed algorithm can be applied for mental workload monitoring.**

Keywords-EEG; Feature Classification; Mental Workload; Multitasking

I. INTRODUCTION

Mental workload and multitasking are topics that have gained increased attention in recent years, given their importance in areas of operator performance and HCI research [1]. In literature, mental workload is commonly defined as the extent to which human mental resource is able to meet the cognitive demands of the task [2]. In defining multitasking, two cases are often considered: sequential multitasking, where the tasks are performed sequentially with obvious task switching and concurrent multitasking, where tasks are performed almost simultaneously [3]. Mental workload and multitasking are closely related, as multitasking is able to induce a high mental workload by increasing the attentional strain on the operator, with the introduction of additional tasks. Mental resources are spent on constantly monitoring and attending the various tasks as required [4]. This can be distinguished from high mental workload induced from a single task of differing difficulty, which seeks to increase mental workload by raising the task complexity. In this case, cognitive resources are dedicated in performing the more complex processing required in a single difficult task [5].

Methods to measure mental workload include subjective methods using participant filled forms [6] and objective methods involving the usage of psychophysiological measurements [7]. It is found that electroencephalography (EEG) provides one of the best methods to measure mental workload, with its high temporal resolution [7]. In the present study, mental workload related to multitasking activity is explored using EEG features and an algorithm for the recognition of the different levels of mental workload is proposed.

II. RELATED WORK

A. Experiments

Mental workload experiments often induce different levels of workload either by increasing the difficulty of a single task or by imposing a multitasking condition. For mental workload induction with a single task, a mental arithmetic task is implemented to classify two distinct brain states of no workload and high workload [8]. Another similar experiment implemented an arithmetic task with seven levels of difficulty and studied the applicability of various EEG measurements in showing a relationship with increasing workload levels [9]. For mental workload induced by multitasking, a general multitasking task battery containing different tasks in sub windows on the same screen is used to study mental workload resulting from multitasking [10]. To study mental workload induced by specific multitasking conditions, a simulation of actual operating conditions are implemented. For example, a simulation of the tasks required by an air traffic controller [11] or aircraft pilot [12] can be designed.

B. Mental Workload Recognition from EEG

Previous research on mental workload with EEG often uses the power spectral density (PSD) of the classical EEG frequency bands as features of interest [13-17]. Other feature types, such as statistical features [8, 18], Higuchi Fractal Dimension (FD) [8], wavelet entropy [5] and Event Related Potentials (ERP) [17] have been proposed as well. However, aside from PSD features, most of the features used in recognizing mental workload have seen little application in classifying the mental workload state related to multitasking condition. For example, in [5] and [8] mental workload was induced with an arithmetic task of varying difficulty while work [17] used an n-back task. Hence, in the present study we shall compare three of the above feature types, the PSD, statistical, and FD features to study their applicability in classification of the mental workload state related to multitasking.

978-1-4673-7218-3/15/$31.00 ©2015 IEEE

III. Materials and Methods

A. SIMKAP Experiment

12 male subjects selected from the university's post graduate population participated in the multitasking experiment. The experimental procedure is described as follows: First, subjects completed a pre-test EEG recording of 1 minute for each of closed eyes and open eyes. The tests begin at this point with EEG recorded throughout the experiment. The first task involved subjects performing a "no-task" condition of 3 minutes open eyes. This is followed by the SIMKAP simultaneous capacity test [19], which involves 3 main components, 2 of which are related to the baseline tasks and the last component is a combination of the baseline tasks in a multitasking scenario. First, the subjects completed an item matching baseline task of 3 trials with different types of items to be matched in each trial: numbers, letters and shapes. The task involves crossing out the appropriate item on the right panel based on items marked out on the left and lasts 3 minutes for each trial. Next, the subjects performed a 3 minutes auditory based problem solving task where questions are asked through the computer speakers and the subject has to answer accordingly by clicking the correct response on the screen. The questions asked can be arithmetic, making comparisons or to identify similar objects. Finally, the subjects perform the multitasking task, shown in Fig. 1, which combines the previous activities and adds a few variations to the auditory task: consult telephone book, consult calendar and answering the question at the required time. The multitasking activity has duration of 18 minutes. After each activity, participants are asked to rate their mental workload level on a scale of 1-9 for the task with a questionnaire, shown in Fig. 2.

From the experiment, we are able to identify 4 possible levels of mental workload and we label them from 1-4 in the following manner: Level 1 – No task condition, Level 2 – Visual matching task, Level 3 – Auditory questions task and Level 4 – Multitasking. We hypothesize that through classification of EEG features, we will be able to distinguish between these 4 levels of mental workload with high accuracy, and rate the cognitive demand required for each task based on pairwise comparison of classification accuracy with the baseline no task condition. Intuitively, one would expect the multitasking condition to require the highest level of cognitive workload. However, between the visual and auditory task, it would be difficult to suggest an intuitive prediction of which task would require more mental resources to perform, although both tasks seem uncomplicated. In the experiment, we used a questionnaire for each subject to rate their mental workload on a 1-9 scale for verification with EEG data. We aim to find out from EEG data, the level of workload for each of these tasks and validate it with ratings from the questionnaire.

EEG signals were recorded from the 14 channels of the Emotiv EPOC device with a sampling frequency of 128Hz and 16 bit A/D resolution. The Emotiv EPOC device used is convenient to set up for recording, and provides comparable performance to a conventional EEG device at an affordable price [20]. The 14 electrode positions used are AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, according to the 10-20 international system [21]. Care was taken in ensuring correct positioning of the device before any recording was performed. During recording, the raw EEG signals were transmitted to the computer through the Bluetooth functionality of the device and data was recorded with the Emotiv TestBench software.

B. EEG Data Processing

Artifact removal was first performed on raw EEG signal with a 2-42 Hz FIR band pass filter using Matlab. Filtered EEG data of 3 minutes duration for each task condition: no-task, matching task, auditory task and multitask were extracted from the continuous EEG signal. As the multitask condition is of 18 minute duration, only the last 3 minutes of EEG data was considered to maintain uniformity with the other tasks. The first and last 15 seconds of EEG data for each 3 minute segment was removed to reduce the effects of unwanted data from any activity in-between tasks. The resulting data for each segment was further divided into 5 partitions of 30 seconds each for feature extraction and cross validation. Power features of the EEG frequency bands related to mental workload: θ band (4-8 Hz), α band (8-12 Hz) and β band (12-30Hz) were extracted with Fourier Transform. Six statistical features, the mean and standard deviation of the filtered signal, mean of the absolute value of the first and second difference of the filtered signal, and mean of the absolute value of the first and second difference of the normalized filtered signal were calculated. FD features were extracted by the Higuchi method [22]. A sliding window of size 512 and shift 128 was used to extract the features. A 5 fold cross validation was performed using Support Vector Machine (SVM) and k Nearest Neighbors (k-NN). Both classifiers have been reported to be reliable for classification of EEG emotion data [23]. In this study, classification was performed for 4 classes and 2 classes with the aim of comparing accuracy between classifiers for mental workload data from EEG. Parameters used for the SVM classifier are: Polynomial kernel with the value of *gamma* set to 1, *coef* set to 1 and order d set to 5. Parameters used for the k-NN classifier are: k set to 1.

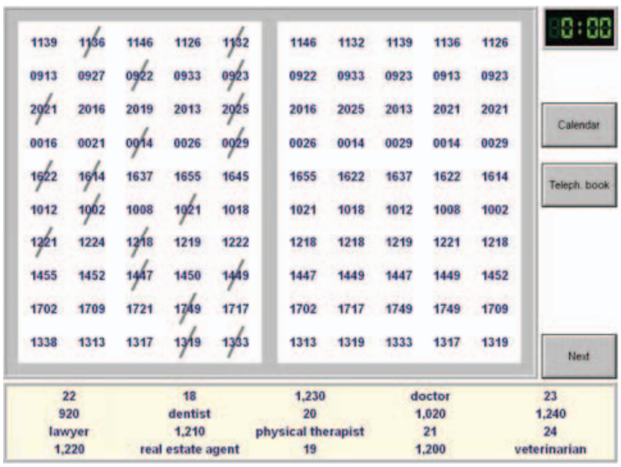

Figure 1. Screenshot of the SIMKAP multitask test. Subjects are to mark items in the right panel by matching those already crossed out on the left panel. Responses to auditory questions are completed by selecting the correct answer from the bottom panel.

Defining **cognitive workload as the amount of mental effort**, on the scale(1-9) below, rate the cognitive challenge involved in the task of this segment with 1 being the lowest and 9 being the highest.

1 2 3 4 5 6 7 8 9

Figure 2. Questionnaire on a 1-9 scale for rating of mental workload, which subjects were required to fill after each task.

IV. RESULTS

The average classification accuracy for 12 subjects is shown in Table 1, using different combination of features: PSD, statistical, FD, PSD & statistical, PSD & FD, statistical & FD, all features. Standard deviation of the average classification accuracy for each feature is shown in the brackets. The best classification accuracy for SVM classifier was obtained using the combination of statistical and FD features, which achieved 80.09% accuracy for 4 classes and 90.39% accuracy for 2 classes with the lowest standard deviation of 4.63. With the k-NN classifier for 4 classes, the accuracy was lower, at 70.28% for 4 classes and 84.97% for 2 classes. For SVM classifier, a one way ANOVA gives a small p value ($p<0.01$), indicating that the proposed combination of statistical and fractal dimension features give significant improvement of accuracy. Fig. 3 ranks the classification accuracy of different feature combinations using SVM classifier compared to k-NN classifier for 4 classes. For all feature combinations, SVM classifier achieved on average a better classification accuracy of 9.56% compared to k-NN classifier.

Table 2 shows the classification accuracy for discerning between different levels 1-4 of mental workload, using different feature combinations for classification. The standard deviation for the average classification accuracy for each pair of level combination is indicated in brackets. The highest average classification accuracy achieved is 94.92% for differentiating between mental workload levels 1 and 4. Fig. 4 ranks the different combinations for 2 levels according to their average classification accuracy. It was found that, compared to the baseline "no task" condition (level 1), multitasking (level 4) has the highest classification accuracy, followed by the visual matching task (level 2) and finally the auditory questions task (level 3). Pairwise classification between levels 2 & 3 and levels 2 & 4 were less accurate, with a similar average classification accuracy of 81.73% and 81.88%.

TABLE I. AVERAGE CLASSIFICATION ACCURACY OF DIFFERENT FEATURE COMBINATIONS FOR 2 AND 4 CLASS SVM AND K-NN CLASSIFIERS

Classifier	PSD	Stat.	FD	PSD + Stat.	PSD + FD	Stat. + FD	All
2 Class SVM	84.16% (6.39)	89.79% (5.22)	82.92% (6.48)	88.05% (4.97)	85.79% (5.80)	90.39% (4.63)	88.47% (4.96)
2 Class k-NN	79.32% (5.24)	84.67% (6.45)	80.27% (6.08)	81.69% (4.95)	80.04% (4.70)	84.97% (6.39)	81.91% (4.82)
4 Class SVM	69.49% (8.03)	78.56% (9.95)	65.26% (10.06)	75.90% (8.18)	71.78% (8.14)	80.09% (8.59)	76.77% (8.37)
4 Class k-NN	60.14% (7.60)	69.65% (10.17)	60.63% (9.05)	64.15% (8.42)	61.47% (6.48)	70.28% (9.97)	64.61% (7.95)

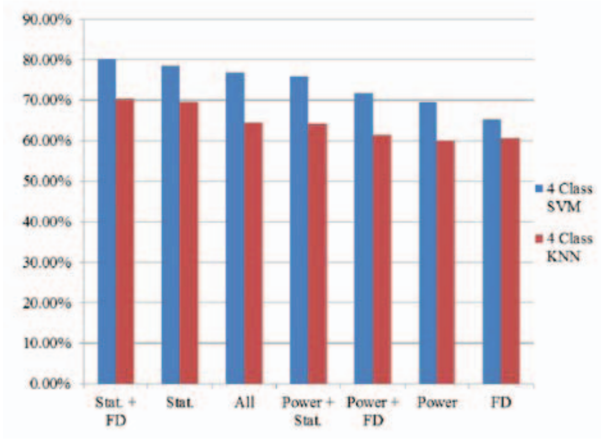

Figure 3. Comparison of classification accuracy for all feature combinations for 4 classes with SVM and k-NN classifiers.

TABLE II. FEATURE CLASSIFICATION ACCURACY 2 LEVELS SVM

Classification Feature(s)	Levels					
	1 & 2	1 & 3	1 & 4	2 & 3	2 & 4	3 & 4
PSD	91.20%	87.41%	91.70%	78.70%	78.46%	77.50%
Stat.	95.06%	92.96%	97.16%	84.14%	85.52%	83.89%
FD	90.28%	86.11%	93.95%	76.27%	76.48%	74.41%
PSD + Stat.	94.01%	91.08%	95.59%	83.70%	83.15%	80.74%
PSD+ FD	92.31%	89.54%	93.11%	80.62%	79.32%	79.81%
Stat. + FD	95.46%	94.29%	97.16%	84.72%	86.63%	84.07%
All	94.20%	92.19%	95.77%	83.98%	83.61%	81.08%
Average	93.22% (1.98)	90.51% (2.98)	94.92% (2.07)	81.73% (3.27)	81.88% (3.83)	80.21% (3.43)

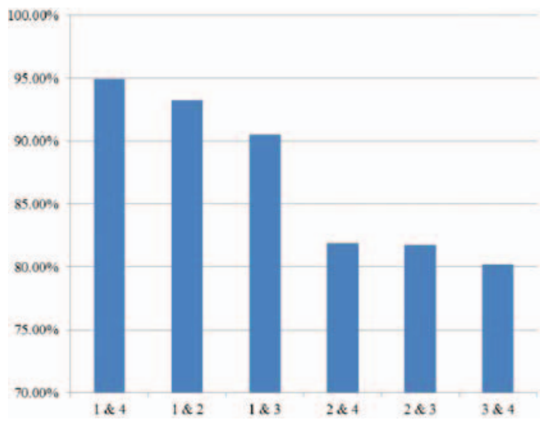

Figure 4. Ranking of classification accuracy between different combinations of 2 mental workload levels

V. DISCUSSION

In this study, we conducted an experiment with 12 male subjects taking the SIMKAP simultaneous capacity test, to classify mental workload related to multitasking. 4 possible levels of mental workload, from level 1 to 4 were identified: no-task, visual matching task, auditory questions task and multitasking. EEG data was recorded from 14 channels of the Emotiv device: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, according to the 10-20 international system, and processing was done to extract the PSD, statistical and FD features from the EEG signal. Classification of mental workload was performed using SVM and k-NN classifiers with fivefold cross validation for 4 classes and 2 classes. The best feature combination was statistical and FD features which achieved 90.39% for 2 classes and 80.09% for 4 classes using SVM. The combination of statistical and FD features significantly outperformed the best reported features including PSD features, with FD feature improving the performance when combined due to its non-linear property. SVM classifier performed better than k-NN classifier with an average higher accuracy of 9.56% for 4 classes, which agrees with findings in [23].

We were able to verify that the EEG data was able to accurately classify the overall mental workload level for each task. This was determined by comparing the classification accuracy of each component task in the SIMKAP test to a baseline "no task" condition, under the assumption that higher classification accuracy would be indicative of a task with higher cognitive workload. From Fig. 4, we found that the multitasking indeed required the highest mental effort, followed by the visual task and the audio task. This was further verified from the questionnaire ranking of the 12 subjects that, on average, the tasks followed the above mentioned ranking in terms of mental workload required.

In this experiment with 12 subjects, we were able to validate EEG data with subjective ratings in distinguishing the level of mental workload required in each task, hence we propose the above protocol which would be able to recognize different levels of task mental workload, along with a feature set comprising of statistical and FD features and SVM classifier for classification.

ACKNOWLEDGMENT

The work is supported by Fraunhofer IDM@NTU, which is funded by the National Research Foundation (NRF) and managed through the multi-agency Interactive & Digital Media Programme Office (IDMPO).

REFERENCES

[1] K. Ryu and R. Myung, "Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic," *International Journal of Industrial Ergonomics,* vol. 35, pp. 991-1009, 2005.

[2] M. S. Young, K. A. Brookhuis, C. D. Wickens, and P. A. Hancock, "State of science: mental workload in ergonomics," *Ergonomics,* pp. 1-17, 2014.

[3] V. Jez, "Searching for the meaning of multitasking," *NOKOBIT,* vol. 2011, 2011.

[4] R. S. Gutzwiller, C. D. Wickens, and B. A. Clegg, "Workload overload modeling An experiment with MATB II to inform a computational model of task management," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2014, pp. 849-853.

[5] P. Zarjam, J. Epps, F. Chen, and N. H. Lovell, "Estimating cognitive workload using wavelet entropy-based features during an arithmetic task," *Computers in biology and medicine,* vol. 43, pp. 2186-2195, 2013.

[6] S. Rubio, E. Díaz, J. Martín, and J. M. Puente, "Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods," *Applied Psychology,* vol. 53, pp. 61-86, 2004.

[7] M. A. Hogervorst, A.-M. Brouwer, and J. B. van Erp, "Combining and comparing EEG, peripheral physiology and eye-related measures for the assessment of mental workload," *Frontiers in neuroscience,* vol. 8, 2014.

[8] Q. Wang and O. Sourina, "Real-time mental arithmetic task recognition from EEG signals," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 21, pp. 225-232, 2013.

[9] P. Zarjam, J. Epps, N. H. Lovell, and F. Chen, "Characterization of memory load in an arithmetic task using non-linear analysis of EEG signals," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 3519-3522.

[10] A. Holm, K. Lukander, J. Korpela, M. Sallinen, and K. M. Müller, "Estimating brain load from the EEG," *The Scientific World Journal,* vol. 9, pp. 639-651, 2009.

[11] M. Z. Weiland, D. M. Roberts, M. S. Fine, and M. S. Caywood, "Real Time Research Methods Monitoring Air Traffic Controller Workload During Simulation Studies Using Electroencephalography (EEG)," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2013, pp. 1615-1619.

[12] R. J. Gentili, J. C. Rietschel, K. J. Jaquess, L.-C. Lo, C. M. Prevost, M. W. Miller*, et al.*, "Brain biomarkers based assessment of cognitive workload in pilots under various task demands," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, 2014, pp. 5860-5863.

[13] J. B. Noel, K. W. Bauer, and J. W. Lanning, "Improving pilot mental workload classification through feature exploitation and combination: a feasibility study," *Computers & operations research,* vol. 32, pp. 2713-2730, 2005.

[14] D. Heger, F. Putze, and T. Schultz, "Online workload recognition from EEG data during cognitive tests and human-machine interaction," in *KI 2010: Advances in Artificial Intelligence*, ed: Springer, 2010, pp. 410-417.

[15] B. Rebsamen, K. Kwok, and T. B. Penney, "Evaluation of cognitive workload from EEG during a mental arithmetic task," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2011, pp. 1342-1345.

[16] Y. Ke, H. Qi, F. He, S. Liu, X. Zhao, P. Zhou*, et al.*, "An EEG-based mental workload estimator trained on working memory task can work well under simulated multi-attribute task," *Frontiers in human neuroscience,* vol. 8, 2014.

[17] A.-M. Brouwer, M. A. Hogervorst, J. B. Van Erp, T. Heffelaar, P. H. Zimmerman, and R. Oostenveld, "Estimating workload using EEG spectral power and ERPs in the n-back task," *Journal of neural engineering,* vol. 9, p. 045008, 2012.

[18] S.-M. Zhou, J. Q. Gan, and F. Sepulveda, "Classifying mental tasks based on features of higher-order statistics from EEG signals in brain–computer interface," *Information Sciences,* vol. 178, pp. 1629-1640, 2008.

[19] O. Bratfisch and E. Hagman, "SIMKAP–Simultankapazität/Multi-Tasking," *Mödling: Schuhfried GmbH,* 2008.

[20] K. Stytsenko, E. Jablonskis, and C. Prahm, "Evaluation of consumer EEG device Emotiv EPOC," in *MEi: CogSci Conference 2011, Ljubljana*, 2011.

[21] R. W. Homan, J. Herman, and P. Purdy, "Cerebral location of international 10–20 system electrode placement," *Electroencephalography and clinical neurophysiology,* vol. 66, pp. 376-382, 1987.

[22] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Physica D: Nonlinear Phenomena,* vol. 31, pp. 277-283, 1988.

[23] A. T. Sohaib, S. Qureshi, J. Hagelbäck, O. Hilborn, and P. Jerčić, "Evaluating classifiers for emotion recognition using EEG," in Foundations of Augmented Cognition, ed: Springer, 2013, pp. 4