Statistical Analysis of Brain Connectivity Estimators During Distracted Driving

Dulan Perera, Yu-Kai Wang, Chin-Teng Lin, *Fellow, IEEE,* Jinchuan Zheng, *Senior Member, IEEE*
Hung T. Nguyen, *Senior Member, IEEE,* Rifai Chai, *Senior Member, IEEE*

***Abstract*— This paper presents comparison of brain connectivity estimators of distracted drivers and non-distracted drivers based on statistical analysis. Twelve healthy volunteers with more than one year of driving experience participated in this experiment. Lane-keeping tasks and the Math problem-solving task were introduced in the experiment and EEGs (electroencephalogram) were used to record the brain waves. Granger geweke causality (GGC), directed transfer function (DTF) and partial directed coherence (PDC) brain connectivity estimation methods were used in brain connectivity analysis. Correlation test and a student's t-test were conducted on the connectivity matrixes. Results show a significant difference between the mean of distracted drivers and non-distracted drivers' brain connectivity matrixes. GGC and DTF methods students t-tests shows a p-value below 0.05 with the correlation coefficients varying from 0.62 to 0.38. PDC connectivity estimation method does not show a significant deference between the connectivity matrixes means unless it is compared with lane keeping task and the normal driving task. Furthermore, it shows a strong uphill correlation between the connectivity matrixes.**

I. Introduction

Driving is an action that requires drivers to direct their full attention while controlling the vehicle [1]. In recent years, distracted driving has become a significant threat to people who commute daily on the roads [2]. Devices or acts that cause the drivers that lead the attention away from driving is identified as distractions. In general, distractions while driving can be divided into four main categories. (i) Visual distraction: looking at something other than the road is called visual distraction, (ii) Auditory distraction: Hearing things which are not related to the driving, (iii) Manual distraction: Engaging in activities which are not related to driving, (iv) Cognitive distraction: Thinking about something other than driving.

Distracted driving can lead to mortality, injuries and property damages. Detection of a distraction while driving could prevent those above-mentioned accidents. Detection of the distraction could be done by monitoring the brain activities. Electroencephalogram (EEG) is a desirable way to monitor brain activities such as distractions and fatigues non-invasively [3] [4] [5]. The organization of the brain and the patterns of links can be expressed by the brain connectivity estimators. Brain connectivity can be subcategorized into structural connectivity, functional connectivity, and effective connectivity [6].

Functional brain connectivity can be either estimated in the time domain or the frequency domain. Granger geweke causality (GGC) connectivity estimation method is one of the methods that can use to identify the directed time-domain brain connectivity. Whereas directed transfer function and the partial directed coherence are the states of the art connectivity analysis methods in the frequency domain.

In this study, Granger geweke causality, directed transfer function (DTF) and partial directed coherence (PDC) is used as the brain connectivity estimation methods [7]. Three main conditions were introduced to investigate brain connectivity. Regular driving through a highway, a lane-keeping task and finally a problem-solving task to introduce a distraction.

This study aims to conclude whether there is a significant difference in brain connectivity estimation methods depending on the distracted driver and a non-distracted driver. Furthermore, linear correlations in brain connectivity estimation methods between a distracted driver and a non-distracted driver were determined.

II. Methodology

A. General Structure

The general structure of the study is shown in Fig 1. In the first stage, distracted driving data is collected from 12 subjects by using 32 EEG channels for 15 minutes of the experiment session [8]. In the second stage, continuous data is down-sampled and filtered out before the event-related EEG epoch extraction. After the required data extraction connectivity analysis is done for each condition in the EEG band from 1Hz to 30 Hz. Statistical analysis is done in the third stage. Correlation tests and student's t-tests are used to analyze the connectivity matrixes between the conditions.

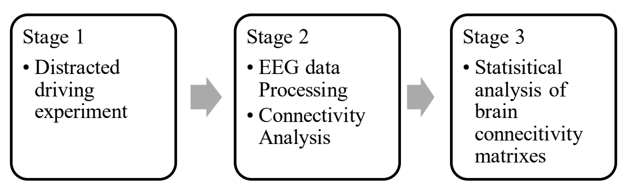

Figure 1. Block diagram of the analysis for the experiment

Figure 2.

B. Data Collection

The data collection experiment was done at the National Chiao Tung University, Hsinchu, Taiwan. Continuous EEG data were collected from 12 healthy participants with more than one year of driving experience with an average age of 24.3 years. Each one of them had normal vision or the corrected normal vision. Drugs, alcohol, and caffeine were forbidden before the experiment. Each participant had two fifteen-minute training sessions before the experiment took place. In these training sessions, participants get to be familiar with the lane-keeping tasks, and problem-solving tasks [5].

Dulan Perera, Rifai Chai, Jinchuan Zheng and Hung T. Nguyen are with Faculty of Science Engineering and Technology, Swinburne University of Technology, Hawthorn, VIC 3122, Australia (e-mail: dgperera@swin.edu.au; rchai@swin.edu.au; jzheng@swin.edu.au hungnguyen@swin.edu.au).

Yu-Kai Wang and Chin-Teng Lin are with School of Computer Science, Faculty of Engineering and Information Technology, University of Technology Sydney, Broadway NSW 2007, Australia (e-mail: Yukai.Wang@uts.edu.au; Chin-Teng.Lin @uts.edu.au)

To get a more realistic EEG data, experiment took place in a dynamic motion simulator with a virtual reality environment, where the car was speeding at a speed of 100Km/hr. in the third lane of a four-lane highway. Two types of experimental conditions occurred randomly though out the fifteen-minute experiment sessions (Lane-keeping task, Math problem-solving task) as shown in Figure 2. math problem-solving task is introduced to manufacture a distraction for the driver

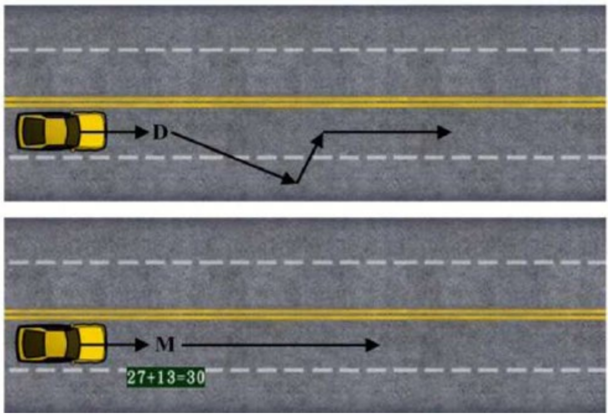

Figure 3. Two conditions in the study. D: lane deviation occurs and M: Math equation appears [5].

C. Experiment Condition

The first condition was the car drifted randomly to either to right or the left side from the current lane and the participant must move it back to the designated lane (Lane 3 from the left). The second condition was a simple Math equation that appears on the screen and the driver must verify whether the math equation is incorrect or correct by pressing the buttons located in the steering wheel. The ratio of the appearance of an incorrect equation and a correct equation was 50:50. The difficulty of the equations remained to be the same throughout the experiment. If the equation appears to be correct, participants were instructed to press the right button on the steering wheel and for the incorrect equation the left button on the steering wheel. The intervals between two conservative trials were 6s to 8s [9].

D. EEG data Acquisition and Preprocessing

EEG data is acquired using a modified 10/20 BCI system with 32 Ag/AgCl electrodes. Data is recorded with a 16-bit quantization level at the frequency of 500Hz. By using a conductive gel impedance was kept under 10kΩ. Raw EEG data and the 32 EEG channels are shown in the Figure 3.The 32 Channels are: FP1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4, P3, Pz, P4, T3, T4, T5, T6, O1, Oz, O2,A1,A2,FT8, FT7,TP7,TP8.

Pre-processing of the EEG data is done by using MATLAB's EEGLab extension. Figure 4 shows the raw input data to the EEGLab Extension [10]. EEG session data collected from each participant was down sampled to 250Hz. 50Hz low pass filter was used to reduce the noise and 0.5Hz high pass filter was used to remove the DC drift. EEG Data for three scenarios, driving data, math problem-solving data and the normal data (EEG data when there is no experiment condition) were extracted from the continuous EEG signal. Reference channels (A1&A2) were removed before the connectivity analysis.

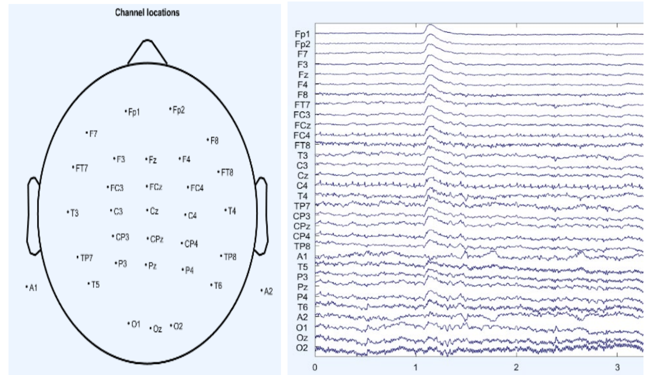

Figure 4. The location and raw data of 26-EEG channels for this study

E. EEG data Acquisition and Preprocessing

Brain connectivity estimators (GGC, DTF, PDC) is defined in the Multivariate autoregressive model framework. AR model can be represented as follows:

$$X(t) = \sum_{j=1}^{p} A(j)X(t-j) + E(t) \tag{1}$$

where sample of data is denoted as X(t) can be expressed in a given time t as sum of previous p samples from the set of k signals weighted model coeeficients A and a random value E.

Sample of data at a time t can be characterized by the sum of previous samples p. Where p is defined as the model order. Model order can be estimated using the Akaike information criterion (AIC), Bayesian information criterion (BIC) (Schwarz-Bayes Criterion (SBC)), Akaike's final prediction error criterion (FPE) and the Hannan-Quinn criterion (HQ).

To determine the causality relations in the time domain Granger geweke causality index was used.

$$GCI_{i\rightarrow j}(t) = \ln\left(\frac{V_{i,n}(t)}{V_{i,n-1}(t)}\right) \tag{2}$$

where residual variance is for n, n-1 dimensional MVAR are denoted as $V_{i,n}(t), V_{i,n-1}(t)$.

Directed transfer function (DTF) which is described in the equation 2 shows the casual influence of the channel j on the channel i.

$$\gamma_{ij}^2(f) = \frac{|H_{ij}(f)|^2}{\sum_{m=1}^{k}|H_{im}(f)|^2} \tag{3}$$

where $H_{ij}(f)$ denotes the elements of the multivariate autoregression model transfer matrix.

Partial directed coherence (PDC) can be used for detection of the directed and cascade flows. PDC can be ditermind as follows:

$$P_{ij}(f) = \frac{A_{ij}(f)}{\sqrt{a_j^*(f)a_j(f)}} \tag{4}$$

where $A_{ij}(f)$ denotes an element from the Fourier transform matrix from the multivariate autoregression model coefficient $A(t)$ and jth colum of $A(f)$ is denoted as a_j.

Pearson's correlation test and student's t-test were conducted as the statistical analysis method to compare the brain connectivity estimation methods on the non-distracted driving and distracted driving. Pearson's correlation coefficient R elaborates the correlation values of the tested scenarios and the p-value from the students t-test elaborates the significant difference between the connectivity matrixes in the given scenarios.

III. RESULTS

A. Source Information Flow Toolbox (SIFT) results

EEGLab's source information flow toolbox was used for the brain connectivity analysis [11]. Output of the AIC, BIC, FPE, HQ from the model order range of 1-30, the optimal model order was selected using the min of the mean curve and elbow of the mean curve methods. Output of the one participant is shown in the Figure 4,5. Model order selection summary is shown in the Table 1.

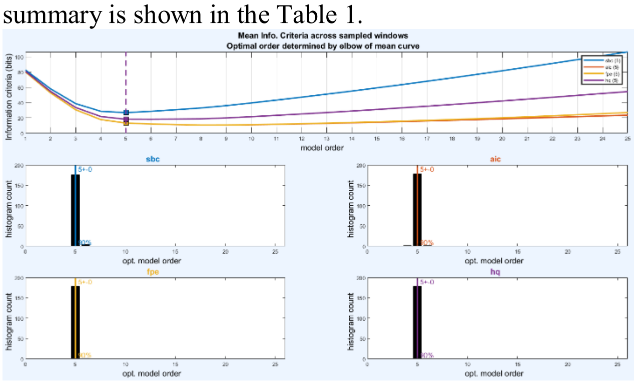

Figure 4: Output of the min of the mean curve method for a subject

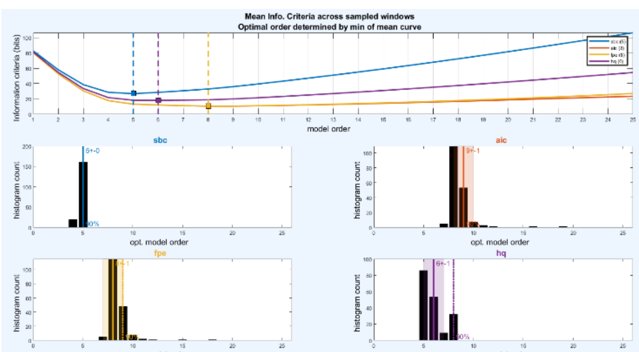

Figure 5: Output of the elbow of the mean curve method for a subject

Table I shows the summary from model order selection from the 12 participants. Four criteria were used to check for the model order. For all three scenarios and four criteria, the elbow of the mean curve method gave the model order as 5 for all 12 participants. The SBC min of the mean curve method gave 5 as the model order for all three scenarios. AIC, FPE with the min of the mean curve model order selection gave 9 for all three scenarios. whereas HQ with the min of the mean curve model order selection gave 8 for all scenarios.

TABLE I: Model Order Selection Summary

Criteria	Lane keeping task		Math problem solving task		Normal driving task	
	Elbow	Min	Elbow	Min	Elbow	Min
SBC	5	5	5	5	5	5
AIC	5	9	5	9	5	9
FPE	5	9	5	9	5	9
HQ	5	8	5	8	5	8

B. Statistical Analysis

Correlation test was conducted to find the linear dependence between two scenarios and student's t-test is used to determine the significant difference between the two scenarios. Student's t-tests and the correlation tests were performed for each participant for the scenarios of Lane-keeping task vs. Math problem-solving task, Lane-keeping task vs. Normal driving, Math-problem solving task vs. Normal driving. After getting the p-values for each participant, the averages of p-values are calculated.

TABLE II: Average p-values from The Student's T-tests

Lane deviation vs. Math problem solving task			Lane deviation vs. Normal driving			Math problem solving task vs. Normal driving		
DTF	GGC	PDC	DTF	GGC	PDC	DTF	GGC	PDC
0.0258	0.0441	0.0742	0.0358	0.0146	0.0278	0.0218	0.0172	0.0622

In the comparison of each above scenario, DTF has lower p-values compare to the others. Granger causality connectivity analysis method has p-values varying from 0.0146 - 0.0441. Whereas p-values of PDC are higher in all three scenarios. From the connectivity analysis methods of DTF and GGC with the p-value below 0.05, it suggests that there is a significant difference between the connectivity matrixes of DTF and GGC in each scenario. Even though by analyzing the p-values of PDC connectivity matrixes, it shows a significant difference in the comparison of Lane deviation vs Normal data scenario, in other scenarios it suggests that there is no significant difference in the connectivity matrixes.

TABLE III: R values from the Pearson correlation coefficient

Lane deviation vs. Math problem solving task			Lane deviation vs. Normal driving			Math problem solving task vs. Normal driving		
DTF	GGC	PDC	DTF	GGC	PDC	DTF	GGC	PDC
0.58	0.43	0.77	0.62	0.39	0.77	0.59	0.38	0.76

It is found that a moderate positive linear relationship is exists in the DTF connectivity matrixes between all three tested scenarios. Granger causality connectivity matrixes indicate a weak positive linear relationship between tested scenarios and PDC connectivity matrixes indicate a strong positive linear relationship between the tested scenarios.

IV. DISCUSSION

A. Lane keeping task vs. Math Problem Solving task

When the directed transfer function and the Granger causality connectivity estimation method is used, we can conclude that there is a significant difference between the means of two connectivity matrixes by using the student's t-test. Furthermore, by analysing the Pearson's correlation coefficient, we can conclude there is a moderate positive linear relationship between DTF connectivity matrixes of the lane deviation and the Math equation. Whereas Granger causality connectivity matrixes indicate a weak positive linear relationship between each matrix. PDC connectivity matrixes show no significant difference in the mean values of the matrixes and a strong positive linear correlation between the two matrixes.

B. Lane deviation vs. Math equation (Lane keeping task vs. Normal Driving task)

By using the student's t-test for all three connectivity estimators (DTF, GGC, PDC) matrixes, we can conclude that there is a significant difference between the connectivity matrixes means. Moreover, it is visible by analyzing the Pearson's correlation coefficient that there is a moderate positive linear relationship between the DTF connectivity matrixes, a weak positive linear relationship between the GGC connectivity matrixes and a strong linear correlation between the PDC connectivity matrixes.

C. Lane deviation vs. Math equation (Lane keeping task vs. Normal Driving task)

The student's t-test shows that there is no significant difference between the DTF and GGC connectivity matrixes means. but it implies that there is a significant difference between the PDC connectivity matrixes means. Furthermore, by analysing the Pearson correlation coefficient we can conclude that there is a moderate positive linear relationship between the DTF connectivity matrixes, a weak positive linear relationship between the GGC connectivity matrixes and a strong linear correlation between the PDC connectivity matrixes.

V. Conclusion

In this study, twelve subjects participated in the driving experiment with two types of experimental conditions. Lane-keeping task and the math problem-solving task are the experimental conditions that occurred randomly throughout the experimental period. 32 electrodes were used to capture the brain waves. Three types of epochs, Lane-keeping task, Math problem-solving task, Normal driving task were extracted from the continuous EEG data. GGC, DTF, and PDC connectivity analyses were performed for each data epoch. By conducting a student's t-test with the p-values <0.05, we can conclude that there is a significant difference between the means of DTF and GGC connectivity matrixes for all three scenarios. but for the PDC connectivity matrixes had a significant difference only when the lane-keep task connectivity matrix and the normal driving data connectivity matrix is compared. Furthermore, by conducting the Pearson's correlation test it is found that there is a weak positive correlation between the GGC connectivity matrixes for all three scenarios, a moderate positive correlation between the DTF connectivity matrixes and a strong positive correlation between the PDC connectivity matrixes. These statistical test results imply that there is a significant difference between the distracted and non-distracted DTF and GGC connectivity matrixes.

VI. References

[1] D. Carr and B. Ott, "The Older Adult Driver With Cognitive Impairment: "It's a Very Frustrating Life"," *The Journal of the American Medical Association ,* p. 303(16), 2010.

[2] K. L. Young, J. Charlton, S. Koppel, R. Grzebieta, A. Williamson, J. E. Woolley and T. Senserrick, "Distraction and older drivers: An emerging problem?," *Journal of the Australasian College of Road Safety,* vol. 29, no. 4, pp. 18-29, 2018.

[3] A. Craig, Y. Tran, N. Wijesuriya and H. Nguyen, "Regional brain wave activity changes associated with fatigue," *Psychophysiology,* vol. 49, no. 4, pp. 574-582, 2012.

[4] R. Chai, G. R. Naik, T. N. Nguyen, S. H. Ling, Y. Tran, A. Craig and H. T. Nguyen, "Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system," *IEEE journal of biomedical and health informatics,* vol. 21, no. 3, pp. 715-724, 2016.

[5] Y. k. Wang, S. a. Chen and C. T. Lin, "An Eeg-based brain-computer interface for dual task driving detection," *Neuricomputing,* vol. 129, pp. 85-93, 2014.

[6] A. Bastos and J. Schoffelen , "Tutorial review of functional connectivity analysis and methods and their interpretational pit falls," *Frontiers in Systems Neuroscience,* vol. 9, p. 175, 2016.

[7] L. A. Baccala´ and K. Sameshima, Methods in Brain Connectivity Inference, CRC Press, 2014.

[8] Y.-K. Wang, T.-P. Jung and C.-T. Lin, "EEG-Based Attention Tracking during distracted driving," *IEEE Transactions on Neural systems and rehabilitation engineering,* pp. 1085 -1094, 2015.

[9] Y.-K. Wang, T.-P. Jung and C.-T. Lin, "Theta and alpha oscillations in attentional interaction during distracted driving," *Frontiers in Behavioral Neuoscience,* p. 3, 2018.

[10] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods 134,* pp. 9-21, 2004.

[11] A. Delorme, T. Mullen, C. Kothe and Z. A. Acar, "EEGLAB, SIFT, NFT, BCILAB, and ERICA: New Tools for Advanced EEG Processing," *Computational Intelligence and Neuroscience,* 2011.