Cognitive Workload Estimation through Lateral Driving Performance

Joonwoo Son and Suwan Park
HumanLAB, Daegu Gyeongbuk Institute of Science & Technology

ABSTRACT

This paper presents an empirical approach for estimating driver's cognitive workload using driving performance, especially lateral control ability through readily available sensors such as lane position and steering wheel angle. To develop a real-time approach for detecting cognitive distraction, radial basis probabilistic neural networks (RBPNN) were applied. Data for training and testing the RBPNN models were collected in a simulator experiment in which fifteen participants drove through a highway and were asked to complete auditory recall tasks. The best performing model could detect cognitive workload at the accuracy rate of 73.3 %. The results demonstrated that the standard deviation of lane position and steering wheel reversal rate can be used to detect driver's cognitive distraction in real time.

INTRODUCTION

Recent technological advances have enabled a wide variety of information systems to be integrated into a vehicle in order to increase safety, productivity, and comfort. However, improperly deployed technology can increase driver's workload and, consequently, degrade safety. Thus identification of a driver's workload and spare capacity is crucial in the design of intelligent vehicles. With this knowledge, the in-vehicle information systems (IVIS) can provide timely and affordable information when the driver has the spare capacity to understand and respond it [1].

Workload refers to the amount of resources that is required to perform a particular task. Two major types of driving workload are visual and cognitive workload [2-3]. Visual workload is straightforward, occurring when drivers look away from the road way; it can be reasonably measured by the duration and frequency of glances away from the road. Unlikely visual workload, cognitive workload is difficult to measure directly because it is essentially internal to the driver. Nevertheless, there have been efforts to measure cognitive workload using subjective measures, physiological measures, eye movement measures, and driving performance measures [4-8]. Among those measures, driving performance measures are known to have limitations compared to others due to small changes according to the cognitive workload, although they are easy and less expensive measures to detect the cognitive workload [8].

This paper presents an empirical approach for estimating driver's cognitive workload using driving performance, especially lateral control ability through readily available sensors such as lane position and steering wheel angle. The results suggest that the lateral vehicle control measures including the standard deviation of lane position and steering wheel reversal rate can be used to detect cognitive workload in real time as inputs of RBPNN models.

DRIVING PERFORMANCE AND COGNITIVE WORKLOAD

Some studies have shown that cognitive distraction undermines driving performance by disrupting the allocation of visual attention to the driving scene and the processing of attended information. Consequently, cognitive workload leads to significantly reduced lane keeping variation and increased response times to sudden obstacles. In this paper, therefore, two driving performance measures, i.e., lateral position variation and steering wheel activity, were selected to assess lateral control ability.

Lateral position variation - Lateral position variation is one of the most commonly used driving behaviors metric. Reduced variation in lateral position when engaged with a cognitive task could be interpreted as a symptom of driver overload and increased risk of incorrect decisions due to being engaged in a distracting task. Lateral position variation is commonly calculated as the standard deviation of lateral position (SDLP). But, SDLP becomes highly correlated to data duration, because the variations in lane position are rather slow. Thus, new lateral position variation measure, modified standard deviation of lateral position (MSDLP), was proposed in the AIDE project [10]. M SDLP is independent of data length, because it is based on high-pass filtering lane position data before standard deviation is calculated. A high pass filter with 0.1 Hz cut off frequency is applied on lane position data. This makes the variation constant after approximately 10 seconds. The filter that was applied resulted in SDLP being uninfluenced by data lengths over the filter time period.

<u>Steering wheel activity</u> - Cognitive secondary tasks yield increased steering activity. The increase is mainly in smaller steering wheel movements, the majority of which are smaller than 1 degree. This often comes with increased gaze concentration towards the road centre and reduced lateral position variance [10]. The steer wheel reversal rate can be used for measuring the increase of smaller steering wheel movements. It is defined as the number, per minute, of steering wheel reversals larger than a certain minimum angular value (so called the gap size).

RADIAL BASIS PROBABILISTIC NEURAL NETWORKS

In this paper, radial basis probabilistic neural networks are applied for estimating driver's cognitive workload using later control measures of driving performance. Radial basis probabilistic neural networks are a kind of radial basis networks which are suitable for classification problems [11]. When an input is presented, the first layer computes distances from the input vector to the training input vectors, and produces a vector whose elements indicate how close the input is to a training input. The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, a compete transfer function on the output of the second layer picks the maximum of these probabilities, and produces a 1 for that class and a 0 for the other classes.

MODEL CONSTRUCTION

DATA SOURCE

<u>Experimental setup</u> - The experiment was conducted in the DGIST fixed-based driving simulator, which incorporated STISIM Drive™ software and a fixed car cab as shown in Figure 1. Graphical updates to the virtual environment were computed using STI SIM Drive™ based upon inputs recorded from the OEM accelerator, brake and steering wheel which were all augmented with tactile force feedback. The virtual roadway was displayed on a 2.5m by 2.5m wall-mounted screen at a resolution of 1024 x 768. Sensory feedback to the driver was also provided through auditory and kinetic channels. Distance, speed, steering, throttle, and braking inputs were captured at a nominal sampling rate of 30 Hz. A display was installed on the screen beside the rear-view mirror to provide information about the elapsed time and the distance remaining in the drive.

<u>Subjects</u> - Subjects were required to meet the following criteria: age between 25-35, drive on average more than twice a week, be in self-reported good health and free from major medical conditions, not take medications for psychiatric disorders, score 25 or greater on the mini mental status exam [12] to establish reasonable cognitive capacity and situational awareness, and have not previously participated in a simulated driving study. The sample consisted of 15 males, who are in the 25-35 age range (M=27.9, SD=3.13).

<u>Cognitive Workload</u> - An auditory delayed digit recall task was used to create periods of cognitive demand at three distinct levels. This form of n-back task requires participants to say out loud the nth stimulus back in a sequence that is presented via audio recording [13]. The lowest level n-back task is the 0-back where the participant is to immediately repeat out loud the last item presented. At the moderate level (1-back), the next-to-last stimuli is to be repeated. At the most difficult level (2-back), the second-to-the-last stimulus is to be repeated. The n-back was administered as a series of 30 second trials consisting of 10 single digit numbers (0-9) presented in a randomized order at an inter-stimulus interval of 2.1 seconds. Each task period consisted of a set of four trials at a defined level of difficulty resulting in demand periods that were each two minutes long.

<u>Procedure</u> - Following informed consent and completion of a pre-experimental questionnaire, participants received 10 minutes of driving experience and adaptation time in the simulator. The simulation was then stopped and participants were trained in the n-back task while remaining seated in the vehicle. N-back training continued until participants met minimum performance criteria. Performance on the n-back was subsequently formally assessed at each of the three demand levels with 2 minute breaks between each level. When the simulation was resumed, participants drove in good weather through 37km of straight highway. Minutes 5 through 7 were used as a single task driving reference (baseline). Thirty seconds later, 18 seconds of instructions introduced the task (0, 1 or 2-back). Each n-back period was 2 minutes in duration (four 30 second trials). Two minute rest/recovery periods were provided before presenting instructions for the next task. Presentation order of the three levels of task difficulty was randomized across participants.

Figure 1. The DGIST Driving Simulator

MODEL CHARACTERISTICS AND TRAINING

Definition of Cognitive Workload - The cognitive workload was classified into two categories, i.e., low workload and high workload by the RBPNN models. In general, the more tasks a driver is conducting at a time, the more resources he/she is consuming and, therefore, the higher workload he/she is bearing [8]. Based on this assumption, the driving performance data in the dual-task period were labeled as high workload and low workload for single tasks period. However, the cognitive capacity required to perform the same tasks varies from person to person. It means that the workload levels induced by n-back tasks might be differ for different drivers. Therefore, the lowest difficulty level, so called 0-back, was omitted in this paper to reduced individual variation. Consequently, single task period (driving only) was categorized as low workload and dual task periods (1-back and 2-back) were considered as high workload.

Input Features - Two driving performance measures, the standard deviation of lane position (MSDLP) and steering wheel reversal rate (SRR), were selected as lateral control ability indices to estimate the driver's cognitive workload in the RBPNN models.

MSDLP was calculated using 0.1 Hz high pass filtered lateral position data. It can be only applied for data sets longer than 10 seconds and not during lane changes.

SRR was calculated by counting the number of steering wheel reversal from the 2Hz low pass filtered steering wheel angle data. For cognitive workload, the reversal angles, more than 0.1 degree of the gap size, were counted.

Summarizing Parameters of Inputs - In this paper, window size was considered as the summarizing parameter for the inputs. Window size denotes the period over which MSDLP and SRR data were averaged. The comparisons of window size could identify the appropriate length of data that can be summarized to reduce the noise of the input data without losing useful information. This paper considered five window sizes: 2, 5, 10, 15 and 30 seconds.

Model Training and Testing - Radial basis probabilistic neural networks (RBPNN) were used to construct the driver's cognitive workload detection models. In this paper, the models were trained using the NEWPNN function in MATLAB.

For training and testing RBPNN models, data of four task segments, which consist of a single task and three dual tasks, were used. A task was divided into multiple segments based on window size. For example, if the model uses 30 seconds window, one task has four segments as shown in Figure 2. In each task, two segments were used for training and the other segments were used for testing. Since the estimator is always evaluated on the data disjoint from the training data, the performance evaluated through the cross validation

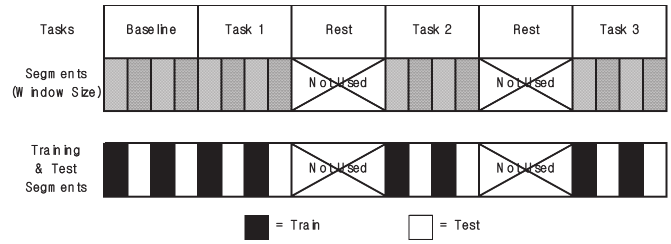

Figure 2. Allocation of Segments to Training and Testing Sets

scheme correctly reflects the actual generalization capability of the derived estimator.

Model performance was evaluated with testing accuracy, which is the ratio of the number of instances correctly identified by the model to the total number of instances in the testing set.

RESULT AND DISCUSSION

According to five time window sizes, workload estimation accuracy rates are described in Table 1. The highest workload estimation accuracy rate in overall model performance was achieved when the time window size was 30 seconds, the longest window size. With 30 seconds window, the accuracy rate of single tasks, i.e., low workload criteria, was 73.3%, and that of dual tasks, i.e., high workload criteria, was 73.3% as well.

From the overall model performance perspective, longer window sizes generated more accurate models, which is consistent with a previous study [8].

Table 1. Model Performance with Different Window Size

Window size (sec)	Workload estimation	Single Task	Dual Task	Total
2	High workload	384	851	-
	Low workload	66	49	-
	Estimation accuracy rate (%)	14.7	94.6	67.9
5	High workload	137	332	-
	Low workload	43	28	-
	Estimation accuracy rate (%)	23.9	92.2	69.4
10	High workload	40	152	-
	Low workload	50	28	-
	Estimation accuracy rate (%)	55.6	84.4	74.8
15	High workload	34	107	-
	Low workload	26	13	-
	Estimation accuracy rate (%)	43.3	89.2	73.9
30	High workload	8	44	-
	Low workload	22	16	-
	Estimation accuracy rate (%)	73.3	73.3	73.3

However, the accuracy rates of single and dual task conditions shows opposite trends with different window size. The longer time window size provides better model performances in the single task detection but poorer performance in the dual task condition.

When the time window size was 15 seconds, for single tasks, the accuracy rate to detect the low workload was 43.3%, and for dual tasks, the accuracy rate to detect the high workload was 89.2%. When the time window size was 10 seconds, for single tasks, the accuracy rate to detect the low workload was 55.6%, and for dual tasks, the accuracy rate to detect the high workload was 84.4%

The results show that the proposed RBPNN models were able to detect driver distraction substantially better than chance performance. Zhang et al. [8] proposed driver cognitive workload estimation and showed the accuracy rate of their method is over 80% through various measures such as driving performance and eye activities by a machine-learning-based DWE (Driver Workload Estimation) development process when the time window size was 30 seconds. The main contributor of the high accuracy rate in their models was eye movement measures, which were obtained from very expensive gaze tracking device. The robustness of eye movement measures is still doubtful because the data were easily influenced by ambient light. However, the proposed method in this paper is easy to implement and compute driver's cognitive workload because it uses only lateral driving performance and there is no need to attach the sensors to a human body like other researches using physiological measures.

CONCLUSION

In this paper, we proposed an empirical approach for estimating driver's cognitive workload using driving performance, especially lateral control ability through readily available sensors such as lane position and steering wheel angle. In order to collect driving data, participants drove through highway in a driving simulator and were asked to complete three different levels of auditory recall tasks. The driver's cognitive workload estimation system was developed using radial basis probabilistic neural network that was implemented by MATLAB and used NEWPNN function.

The results demonstrated that the proposed RBPNN models were able to detect driver distraction substantially better than chance performance, and the standard deviation of lane position and steering wheel reversal rate can be used to detect driver's cognitive distraction in real time. For model parameter selection, longer window sizes generated more accurate models, which is consistent with a previous study [8].

The model performance was assessed with the cross-validation scheme, which is widely adopted by the machine learning community. As a result, the highest workload estimation accuracy rate in overall model performance was 73.3%. Although 73.3% of accuracy is not enough to use in everyday monitoring, it is challenging because the result was achieved from two driving performance measures; MSDLP and SRR. They are easy to collect the data through readily available sensors, and need not to attach additional sensors to human body. And it is also expected that the accuracy can be improved by applying more sophisticated algorithms.

ACKNOWLEDGMENTS

This research was supported in part by Daegu Gyeongbuk Institute of Science and Technology (DGIST) Research Program of the Ministry of Education, Science, and Technology (MEST), and Industrial Strategic Technology Program of the Ministry of Knowledge Economy (MKE).

REFERENCES

1. Son, J., Lee, Y., and Kim, M. (2010), "Impact of traffic environment and cognitive workload on older drivers' behavior in simulated driving", International Journal of Precision Engineering and Manufacturing, Vol. 12, No. 1, pp.135-141.
2. O'Donnel, R.D. and Eggemeier, F.T.(1986), "Workload assessment methodology," in Handbook of Perception and Human Performance, Vol II, Cognitive Processes and Performance, K.R. Boff, L. Kaufman, and J.P. Thomas, Eds. Wiley, New York.
3. Angell, L., Auflick, J., Austria, P.A., Kochhar, D., Tijerina, L., Biever, W., Diptiman, T., Hogsett, J., and Kiger, S. (2006), Driver Workload Metrics Task 2 Final Report, DOT HS 810 635, National Highway Traffic Safety Administration, Washington, DC.
4. Mehler, B., Reimer, B., and Coughlin, J. (2010), Physiological Reactivity to Graded Levels of Cognitive Workload across Three Age Groups: An On-Road Evaluation. Proceedings of the 54th Annual Meeting of the Human Factors and Ergonomics Society, San Francisco.
5. Lenneman, J.K. & Backs, R. W. (2009). Cardiac autonomic control during simulated driving with a concurrent verbal working memory task. Human Factors, Vol. 51, No. 3, pp.404-418.
6. Mehler, B., Reimer, B., Coughlin, J.F., & Dusek, J.A. (2009). The impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. Transportation Research Record, 2138, p. 6-12.
7. Reimer, B. (2009), "Impact of Cognitive Task Complexity on Drivers' Visual Tunneling", Journal of the Transportation Research Board of the National Academies, Vol. 2138, pp. 13-19.
8. Zhang, Y., Owechko, Y. and Zhang J. (2004), "Driver cognitive workload estimation: A data-driven perspective", Proc. IEEE Intell. Transp. Syst. Conf., pp.642.
9. Engström, J., Johansson, E. and Östlund, J. (2005), Effects of visual and cognitive load in real and

simulated motorway driving, Transportation Research Part F: Traffic Psychology and Behaviour 8, pp. 97–120.

10. Östlund, J., Peters, B., Thorslund, B., Engström, J., Markkula, G., Keinath, A., Horst, D., Juch, S., Mattes, S. and Foehl, U. (2004), Driving performance assessment – methods and metrics, Information Society Technologies, No. IST-1-507 674-IP, pp. 118-129.
11. Demuth, H. and Beale, M. (2000), Neural Network Toolbox User's Guide, MathWorks, Inc.
12. Folstein, M.F., Folstein, S.E., and McHugh, P.R. (1975) "Mini-mental state: A practical method for grading the cognitive state of patients for the clinician", Journal of Psychiatric Research, Vol.12, No.3, pp. 189-198.
13. Son, J., Reimer, B., Mehler, B., Pohlmeyer, A. E., Godfrey, K.M., Orszulak, J., Long, J., Kim, M.H., Lee, Y.T. and Coughlin, J. F. (2010). Age and cross-cultural comparison of drivers' cognitive workload and performance in simulated urban driving. International Journal of Automotive Technology, Vol. 11, No. 4, pp.533-539.
14. Mehler, B., Reimer, B., and Coughlin, J.F. (2010). Physiological reactivity to graded levels of cognitive workload across three age groups: An on-road evaluation. Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting, San Francisco, CA, pp.2062-2066.

CONTACT

Joonwoo Son

Mail: json@dgist.ac.kr (joonwooson@gmail.com)

Highest degree: PhD, Mechanical Engineering, Pusan National University

Position title: Principal Investigator, HumanLAB, DGIST

Areas of interest: Design guidelines for older drivers, Driver's workload assessment methodology (On-road & simulator study), Advanced Human-Vehicle Interface, Assess driving style and fuel efficiency, Eco-Drive assistant system