Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Mental Workloads Can Be Objectively Quantified in Real-time Using VOR (Vestibulo-Ocular Reflex)

Goro Obinata*. Satoru Tokuda**
Naoki Shibata***

/* Nagoya University, Aichi, Japan,*
(Tel: +81-52-789-5030; e-mail: obinata@mech.nagoya-u.ac.jp).
*** Wichita State University, Wichita, KS 67208, USA*
(e-mail:sxtokuda@wichita.edu)

Abstract: This present study offers a new method to quantify mental workloads (MWL) utilizing vestibule-ocular reflex (VOR). The VOR method makes use of the relation between a person's VOR responses and his/her mental demands; Human VOR responses can be accurately predicted in a dynamical equation that is a function of the person's head movements, unless the person is engaging in a higher cognitive activity. In this present study, the coherence between the predicted VOR and the observed VOR was as high as 0.92 when there was no additional mental demands. However, the manipulation of MWL in five different tasks (i.e. the n-back task) revealed that the VOR coherences declined with the heavier MWL demands. This shows that MWL can be objectively quantified by measuring the gap between observed VOR responses and the mathematical-model-predicting VOR. This may be applicable in the future to quantifying a vehicle driver's MWL in real-time.

1. INTRODUCTION

Mental workload (MWL) is an essential concept in most Human-Machine systems. Mental workload can be defined as the currently used amount of cognitive resources in a person at a given point in time. Since cognitive resources in humans are limited, human performance is easily deteriorated by heavy MWL.

Many methods have been developed to quantify MWLs. These techniques can be categorized in three groups (Stanton, et al., 2005); (a) primary/secondary task measures, (b) subjective-rating techniques including NASA-TLX, and (c) physiological measures. The third one, physiological measures, is supposed to be the best one to quantifiably measure vehicle driver's MWL because they are objective, less interfering with the driving task and can be measured in real-time.

Vestibule-ocular reflex (VOR), one of the physiological measures, has grabbed the attention of recent researchers and has been examined to assess its effectiveness in quantifying MWLs (e.g. Furman, et al., 2003; Shibata, Obinata & Kajiwara, 2006; and Shibata, Obinata, Kodera & Hamada, 2006). This VOR method of quantifying MWL has at least six major advantages over other existing MWL measures: The VOR method is (1) objective not relying on the driver's subjective ratings, (2) does not interrupt the main tasks, (3) measurable in real-time, (4) accurate, (5) not physically obtrusive, and (6) does not require large equipment. Taking from previous studies, our present study used the n-back tasks as the mental workload demands to determine if VOR responses could be reasonable measures for quantifying a person's MWLs.

2. MODEL OF EYE MOVEMENT

2.1 *Model of Vestibulo-Ocular Reflex (VOR)*

Vestibule-ocular reflex (VOR) is an involuntary eye movement in humans performed to keep an object at a fixed gaze in order to offset their head movements. For example, when you are fixing your gaze at an unmoved object right in front of you, and your head moves downward, you involuntarily move your eyes upward (the opposite direction of the head movement) so that the eyes can keep the fixation on the object.

This present study integrated two parts of VOR models in Fig.1 to calculate the participants' VOR responses: Fig. 1 (I) made by Merfeld and Zupan (2002) and Fig.1 (II) made by Robinson (1981). A mechanism of a VOR model consists of three stages, as shown in Fig.1(I); (1) physical world & sensors, (2) internal processing, and (3) eye movements.

In the first stage, the person's VOR mechanism senses six pieces of head movement information; linear accelerations in three dimensions (α in Fig.1.I) and angular velocities in three dimensions (ω in Fig.1.I). In our study, a magnetic field sensor collected these data.

In the second stage, the person's VOR mechanism integrates these six inputs and calculates appropriate eye movements so that the eyes can follow a visual object when the head is moving. Since this calculation is not the same across people, personalization of the VOR model is required for an accurate simulation. This is explained more in the next section: "Model Identification."

978-1-1234-7890-2/08/$20.00 © 2008 IFAC 10.3182/20080706-5-KR-1001.1635

In the third stage, the person's VOR mechanism executes the calculated eye movements. The eyes can theoretically move on three axes: horizontal, vertical, and torsional rotations. However, this present study utilized only the two main rotations: horizontal and vertical. Our study examined the discrepancies between the model-calculated eye movements and the observed eye movements on these two axes.

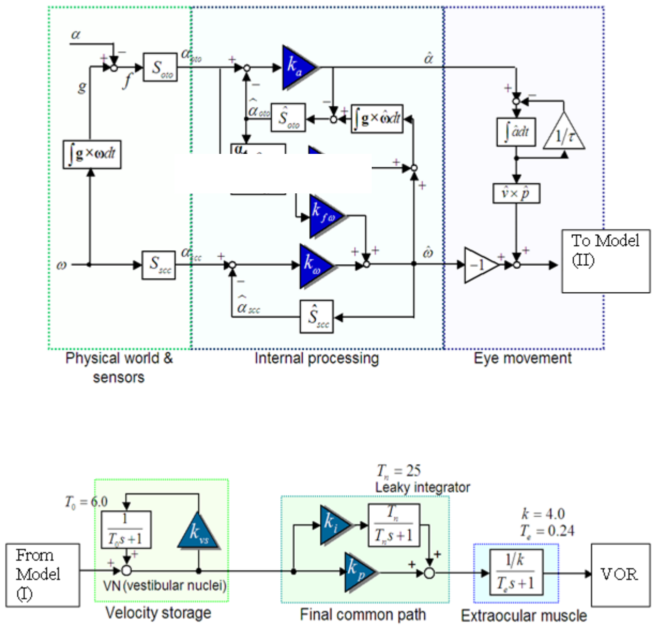

Fig. 1. Model of vestibulo-ocular reflex.

Although we know there is a relationship between the vestibular system and the VOR eye movement, as well as the fact that heavy MWL influences the VOR eye movement, we don't know the cognitive mechanism behind why heavy MWL influences the VOR eye movement. One possible cognitive mechanism is shown in Fig.2, adopted from Honrubia (1997). This model stipulates that common central processors integrate all the sensory inputs and determine the eye movements, including the VOR. These central processors may require some cognitive resources and may therefore be affected by heavy MWL. However, there have been no studies to prove this mechanism. The focus of our study is not on the mechanism behind it, but on the application of how useful the VOR disturbance is to quantify a person's MWL.

2.2 Model Identification

A person's VOR responses are almost perfectly predicted when all the 4 conditions below are satisfied: (1) The person's head is being shaken with frequencies from 1 to 6 hertz; (2) The head movements are quantified and used as the inputs of the VOR model; (3) The VOR's internal processing stage is personalized for the person; and (4) The person is not engaging in a cognitively demanding task.

With the four conditions satisfied, as done in our study, the VOR model predicted the person's VOR responses as perfectly as 0.92 in the coherence between the simulated VOR and the observed VOR.

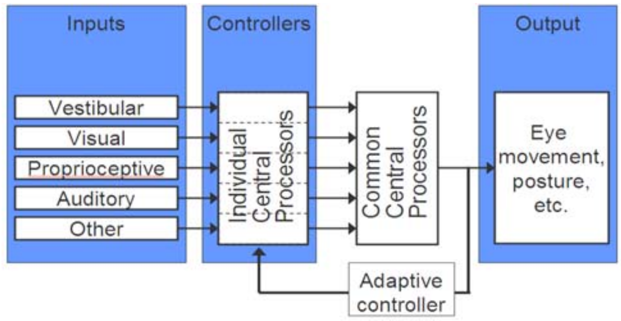

Fig. 2. One possible cognitive mechanism that explains why heavy MWL influences the VOR eye movement. Adopted from Honrubia (1997).

Model identification is the process to identify individual's eye movement dynamics that are represented by the 12 constant parameters shown with the seven triangle shapes in Fig.1. Since our experimental devices did not have the means to measure the torsional eye movement, we used the 8 parameters (Ka, Kf, Kfw, Kw, Kr_vertical, Kr_horizontal, Kv_vertical, and Kv_horizontal). These parameters are constants and are different by person and possibly by time.

This present study used a computer-controlled-shaking chair called "Joy Chair" to induce participant's head movements, which caused the participant's VOR responses. The Joy Chair was set to produce tremors in frequencies from 1 to 6 hertz. Since VOR responses are different by head movement and by individual, our experiment first developed the person's VOR mathematical model by making the computer program observe the person's specific head movement dynamics for 10 seconds while the person stared at a fixed point on a projector screen. We call this stage "the VOR model identification." Fig.3 shows the experimental setup and devices.

The Matlab program, a computer program specialized to identify a mathematical model, analyzed the relation of the head movements with the person's VOR responses and produced a best-fit mathematical model to predict the person's VOR responses in a given situation, considering temporal sequences of both the head movements and the eyeball movements. A best-fit model means that the mathematical equation sets certain constant numbers specific for the person's VOR responses, so that errors between the mathematical model-predicted VOR and the actual VOR responses are minimal.

Eight constant parameters need to be identified to accurately predict the person's specific VOR responses in the Merfeld-Robinson model of Fig.1. Since there are numerous possible combinations of eight constant numbers, it would take forever to find a best-fit mathematical model for each individual participant if all the possible combinations were evaluated. Our study took two shortcuts: a Genetic Algorithm (GA) and the Gradient method of local search in order to moderately quickly find a seemingly best-fit model with certain constant parameters so that the errors were minimized based on the 10-second data for the VOR model identification. For the global search, the GA started with one specific DNA combination of the eight parameters such as -2, 1, 1, 4, 1, 1, 4, and 4. The GA then mutated the DNAs, paired with another DNA, and produced descendant DNAs. The Gradient method of local search determined what tendency of alteration in these eight parameters leads to the minimum error between the calculated VOR and the observed VOR. The combination of these two methods above found the eight parameters that yielded the least error values between the calculated VOR and the observed VOR.

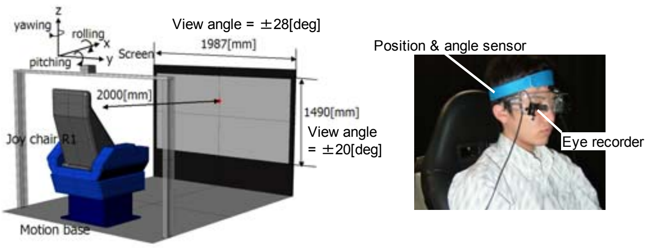

Fig. 3. Experimental setup and devices.

Using these two methods, the goodness-of-fit index to estimate the constant parameters is expressed by

$$J = \sum_{i=1}^{N} \{\theta_{obs}(i) - \theta_{mdl}(i)\}^2 \qquad (1)$$

where θ_{obs} is a time series of observed eye movement data and θ_{mdl} is a time series of the model-estimated eye movement values.

The first step of the model identification in our study was to clean up the raw data. After collecting raw data of eye movements, rapid eye movements such as blinking and saccades were removed from the data set because VOR is independent from these rapid eye movements, and they would have made our VOR analysis difficult. The cleaned-up data was processed by a Matlab program to identify a mathematical model for each participant's VOR responses.

2.3 Identification Results

Fig.4 shows an example of the identification results in several aspects. Fig.4(a) shows the time course of the data for a period of 10 seconds. Fig.4(b) shows the results in the frequency aspect.

Fig.4(a)(I) and (b)(I) show the results before the model identification, using the previously identified general constant parameters, offered by Merfeld and Zupan (2002). Fig.4(a)(II) and (b)(II) shows the results after the model identification, using the constants parameters that were personalized for each participant in our experiment.

Each of the Figures 4(b) has two sets of graphs: horizontal and vertical eye movements. For each of these dimensions of eye movements there are two curved lines: the measured output in the observation (red line) and the simulated output of a mathematical model (blue line).

One of the major advantages of using the VOR as MWL measures is that the VOR can be quantified in real-time. However, in our study, the simulated outputs were calculated after the experiments. The present article uses the word "predict" to mean "to calculate the person's simulated VOR responses using only the past and present information" in the dynamical function of both head movements and personalized constant parameters. Usui et al. (2007) use the VOR measure to quantify MWL in real-time and examine the validity of the VOR measure in real-time.

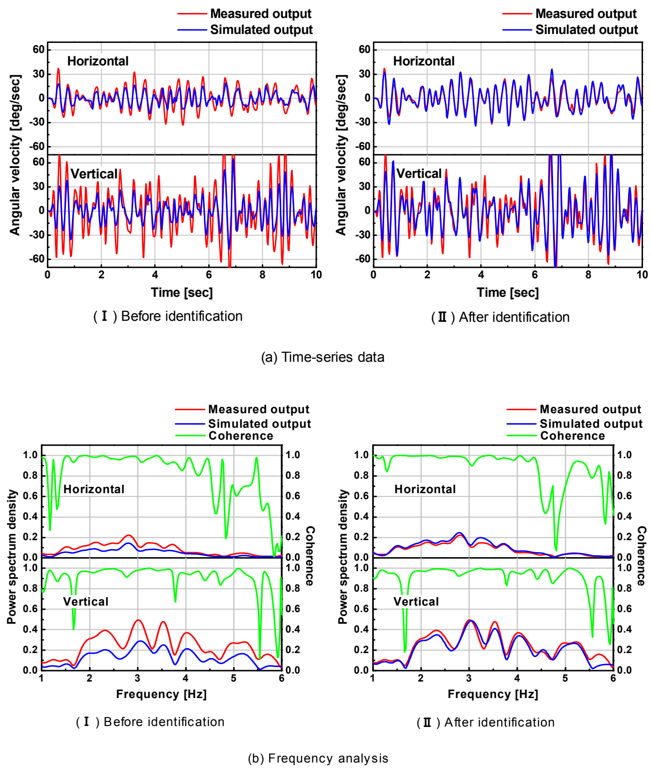

Fig.4. Example of identification result.

As mentioned in the previous section "2.2 Model Identification", the eight constant parameters were personalized for each participant after the model identification. This process of model identification should increase the degree of coherence between the two lines.

Discrepancies between the two lines (blue and red lines in Fig.4.a and 4.b) indicate a low coherence (low in the green lines) at that moment. As seen in Fig.4, the accuracy of the mathematical model to predict the person's eye movements in the time course was higher after the model identification in Fig.4(a)(II) on both horizontal and vertical eye movements than before personalizing the model for the participant in Fig.4(a)(I).

Likewise, Fig.4(b) shows the higher accuracy after the model identification in Fig.4(b)(II) than before in Fig.4(b)(I) in the frequency analysis of the eye movements. VOR responses are usually observed in the frequencies between 1 and 6 hertz of head movements. After the personalized identification in our study, the identified model almost perfectly predicted the actual eye movements in Fig.4(b)(II).

The apparent evidence of overlapping in the graphs is also represented by one measure: the coherent coefficient, which is defined as

$$\gamma^2_{xy}(f) = \frac{|C_{XY}(f)|^2}{P_{XX}(f)P_{YY}(f)} \tag{2}$$

where $P_{XX}(f)$, $P_{YY}(f)$ are power spectra (the two output curves) of the measured eye movement and the predicted signal, respectively, and where $C_{XY}(f)$ is cross spectrum between them. The coherence (drawn with the green lines) takes the unit value of 1.00 when the relation of two signals can be described by a linear differential equation.

Fig.4(b)(II) shows that after the identification of the person's VOR response patterns, the coherence drawn with the green lines are close to the perfect coherence of 1.00 at the majority of the frequency range between 1 and 6 hertz. Since "after identification" in Fig.4(b)(II) has higher coherences than "before identification" in Fig.4(b)(I), our identification seemed to succeed in adjusting the VOR model to the individual participants.

However, among the three participants in our present study, the averages of the model error varied by person from 3 to 6 degrees in eye movement angles out of the experimental eye movement range of 20 degrees on the horizontal axis and 30 degrees on the vertical axis. This indicates that individual differences might have a large impact on VOR measures, although the present study did not focus on the individual differences.

3. METHOD OF EXPERIMENT

3.1 Experiment Procedure

Three male students, aged between 20 and 24, participated in this study. Since the main purpose of this study was to examine if deviation of observed vestibule-ocular reflex (VOR) responses from the modeled VOR was correlated with required mental workloads (MWL), each participant performed in four different MWL conditions (1 controlled and 3 experimental conditions) while the computer-controlled-shaking Joy Chair was causing the participant's head to shake at the frequencies of 1 to 6 hertz. The participants were asked to gaze upon a certain fixed point on a projector screen for 30 seconds on each trial. Each of the 4 conditions was repeated three times, meaning a total of 12 trials were performed by each participant. On each trial, eye movements and reaction time to every verbal presentation were recorded.

The 4 conditions were as follows: (A) a control condition that was the Simple Reaction Task (SRT) condition where the participant was asked to simply hit the button when he heard another alphabet letter, which was provided every 2.5 seconds, (B) 1-back task, (C) 2-back task, and (D) 3-back task.

3.2 The n-back Tasks

The n-back tasks including 1-, 2-, and 3-back tasks impose different amounts of MWL on the person so that experimenters can manipulate the participant's MWL. In our n-back tasks, one alphabet was verbally presented to the participant every 2.5 seconds for 30 seconds on one trial. The participants were asked to hit the yes button when the same letter reappeared after n-events of verbal alphabet presentation and the no button in other cases. The participants were notified of n beforehand. In this present study, n was either 1, 2, or 3.

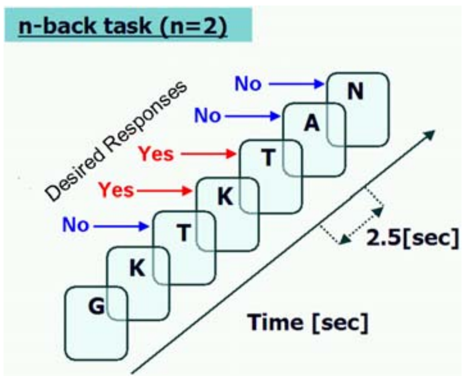

Fig. 5. Verbal n-back task in case of n = 2.

The n-back tasks are usually used in working memory research because the n-back tasks require the person to maintain and update information at a certain pace. Past research has revealed that during the n-back tasks, the activated areas were the frontal association area, temporal association area, and Broca's area (Braver, et al., 1997; Cohen, et al., 1994). In this present study, the words "mental workload" (MWL) and "working memory" are interchangeably used unless specifically indicated. Loads of

MWL are different by person. However, for any person, a higher-number-back task universally requires more MWL than a lower-number-back task.

4. RESULTS

Overall, the results show what we wanted to show; the more demanding mental workload (MWL) tasks induced more discrepancies between the simulated vestibulo-ocular reflex (VOR) responses and the observed VOR responses.

Fig.6 shows the average results of the three participants on the Proportion Correct (PC, the rate of the right answer) and the reaction time in the n-back tasks. The results show that the participants took a longer time to answer in the more demanding n-back tasks, such as the 3-back task, and seemed to try to correctly answer by taking their time in all the three n-back conditions, rather than to simply respond without thinking. This implies that our tasks seemed to appropriately manipulate and impose different MWL levels with the different n-back tasks.

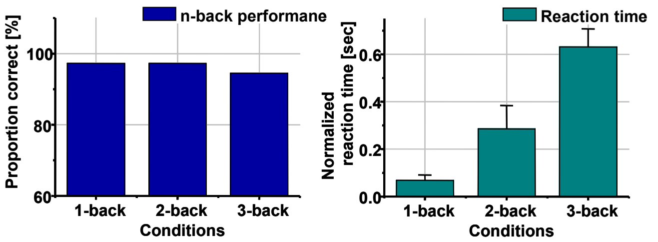

Fig. 6. Proportion correct in percentage and reaction time in the n-back tasks.

Fig.7 shows the results of the power spectrum (red and blue lines) and the coherence (green lines) between the measured VOR and the mathematically simulated VOR in the four conditions: the control condition (Simple Reaction Task), 1-, 2-, and 3-back conditions. The coherence coefficients are indicated in the green lines. However, the frequencies above 4 hertz in the horizontal eye movements were shaded lightly in the graphs because they seemed to have too much noise and be at the error levels, not reflecting the VOR responses much.

The comparisons between the measured VOR and the simulated VOR in the four conditions in Fig.7 indicate several interesting things.

First, there does not seem to be clear distinctions between the power spectrum curves (labeled as Measured and Simulated outputs) to distinguish the 4 conditions in Fig.7(I),(II),(III), and (IV). In the vertical eye movements, all of them had the highest power spectrum at the frequencies between 2.6 and 4.6 hertz. The horizontal axis recorded much lower power spectrum than the vertical axis in all the four conditions. This indicates that just looking at measured VOR and simulated VOR does not tell us the person's MWL.

Second, conversely, the coherence curves (green lines in Fig.7) had relatively distinct characteristics from each other among the four conditions. The coherence curves were highest on both horizontal and vertical axes in the control condition (Fig.7.I), followed by the 1-back task (II). The tasks with the demanding mental workloads such as 2-back (III) and 3-back (IV) tasks had the lower coherence lines.

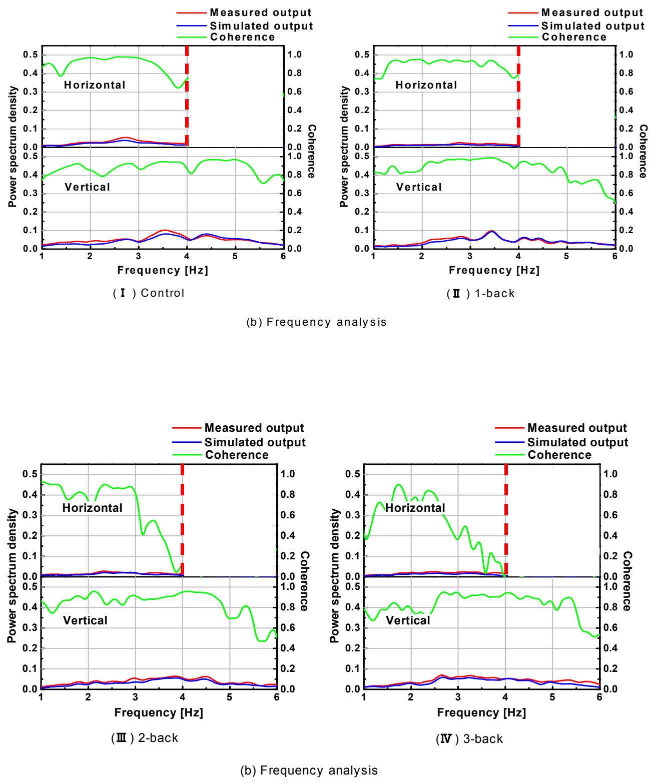

Fig. 7. Power spectrum curves and the coherence in the horizontal and vertical axes in four conditions.

This fact is re-stated again in Fig.8 after simplifying the graphs in one aspect. Fig.8 shows the average of the coherence in the n-back tasks over all the three participants. In addition to SRT and the three n-back tasks, this figure adds the VOR results, in which condition the participants were asked just to stare at a fixed point on the projector screen without doing any additional task.

As expected, the VOR condition achieved the highest coherence, as high as 0.92, closest to the perfect coherence 1.00 among the 5 conditions. Also as expected, the Simple Reaction Task with the coherence of 0.90 followed the VOR condition since there was not much cognitive workload involving this simple reaction task. In regard to the n-back tasks, as the n increased, the supposed mental demand increased, and the deviation of the coherence increased. The coherences were 0.87 for the 1-back task, 0.83 for the 2-back tasks, and 0.80 for the 3-back task. More discrepancies from the baseline value of 0.92 were observed as the supposed MWL increased.

The decrease of the coherence with demanding mental tasks implies that the person's cognitive activities may have interfered with the VOR mechanism in some ways. Our results were consistent with the previous studies by other researchers (e.g. Furman, et al., 2003; Talkwski, et al., 2005; and Yardley, et al., 1999). These results suggest that the VOR measures enable quantification of a MWL.

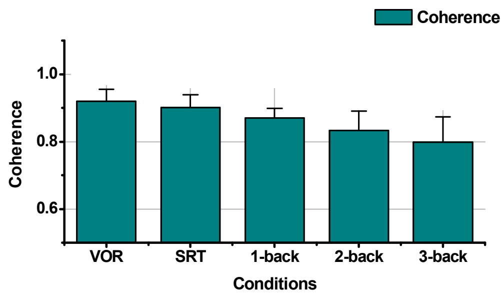

Fig.8. The coherence comparison among 5 situations: VOR with no additional task; Simple Reaction Task to hit a button without judging; and n-back tasks with mentally demanding judgment. Mental demands were correlated with the deviation of the coherence.

5. CONCLUSIONS

This study showed that the vestibule-ocular reflex (VOR) measures can be used as a method to objectively quantify a person's mental workload (MWL).

The results show that the five different levels of mental demands were negatively correlated with the coherence between the simulated VOR responses and the observed VOR responses. When the participants were not engaging in a cognitive task, the VOR responses were predicted with the high coherence of 0.92. However, the coherence was as low as 0.80 when the participants were doing the 3-back task, the most mentally demanding task in our experiments. The observed VOR discrepancies from the simulated VOR responses were probably the result of interference with the human VOR mechanism. This indicates that when unusual VOR responses are observed, the person is most likely to be heavily using his cognitive systems.

Possible applications of this VOR method are many since this method can objectively quantify a person's MWL in real-time without using large devices. Another researcher team in our laboratory is examining the real-time aspect of VOR to answer how accurately the person's MWL can be quantified using VOR measures in real-time.

In future research, our laboratory is planning to examine VOR responses when MWL is varying by seconds. This will reveal how quickly the VOR measures can reflect the person's MWL. We hope someday this technique will detect over-demanding mental conditions of drivers and help to reduce vehicle-involved accidents.

ACKNOWLEDGMENT

We thank Toyota motor Co. Ltd. for their giving us the motivation of this study.