

# Quantitative Evaluation of Mental Workload by Using Model of Involuntary Eye Movement

Goro Obinata<sup>1</sup>, Satoru Tokuda<sup>2</sup>, Katsuyuki Fukuda<sup>1</sup>, and Hiroto Hamada<sup>3</sup>

<sup>1</sup> Nagoya University, EcoTopia Science Institute, Furo-cho, Chikusa-ku,  
464-8603 Nagoya, Japan

<sup>2</sup> Wichita State University, Wichita, Kansas, USA

<sup>3</sup> Toyota Motor Co. Ltd., 1 Toyota-cho, Toyota 471-8572, Japan  
obinata@mech.nagoya-u.ac.jp

**Abstract.** This study considers a new method to quantify mental workloads (MWL) by using mathematical models for reflex movement of eye. Several mathematical models of reflex movements have been proposed and experimentally verified by physiologists. In those models, some models of vestibulo-ocular reflex (VOR) have sufficient accuracy to predict eye movements of individuals. The engagement of brain function to VOR is known in the learning process or in the adaptation process. This leads to the assumption that mental workloads appears in the change of characteristics of VOR. So as to confirm the assumption, we have designed an experimental setup and carried out several experiments. In the experiments, subject's VOR responses have been accurately predicted by the mathematical model which is a dynamical model with the input of head movements and the output of eye movements. The model dynamics have changed while the subject was engaging in a higher cognitive activity. The coherence between the predicted VOR from the identified model of the particular subject and the observed VOR was as high as 0.92 when there was no additional mental demands. However, the changing MWL over five different n-back tasks revealed the clear correlation between the predicted VOR coherences and the MWL demands. This shows that MWL can be objectively quantified by measuring the error between observed VOR responses and the predicted VOR ones from the identified model.

## 1 Introduction

Mental workload (MWL) is an essential concept in most Human-Machine systems. MWL 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 proposed to quantify MWLs. These techniques can be categorized in three groups [1]; (a) primary/secondary task measures, (b) subjective-rating techniques including NASA-TLX, and (c) physiological measures. Unfortunately, most existing MWL measures to date share some of the major disadvantages, especially when measuring vehicle driver's MWL. First, the primary/secondary task

measures cannot quantify a MWL in real-time because it requires a person to perform two different tasks at two different times and compares the performances. Second, subjective-rating techniques require the person to rate the difficulty level of a task usually after a section of a task. This technique does not measure the person's MWL in real-time, does not objectively quantify MWLs, and interrupts the main task to record rating. Third, although physiological measures are good at objectively quantifying mental workload in real-time with no or little interruption of main tasks, most of these physiological measures have other disadvantages. For example, heart rate and respiratory rate measures are not accurate in quantifying MWL. Brain imaging techniques such as fMRI and PET techniques are physically obtrusive to a person in a practical situation. All the previously existing MWL measures have some shortcomings when we use one of those methods to quantify human's MWLs in practical situations. Recently, many researcher groups have started focusing on studying other kinds of physiological measures, such as involuntary reflexes, in the hopes of finding better MWL assessment techniques. For example, vestibule-ocular reflex (VOR) has grabbed the attention of researchers and has been examined to assess its effectiveness in quantifying MWLs [2], [3]. This VOR method of quantifying MWL has six major advantages over other existing MWL measures: The VOR method is (1) objective, (2) does not interrupt the main tasks, (3) measurable in real-time, (4) quickly reflects MWL, (5) not physically obtrusive, and (6) does not require large equipment. Taking from previous studies, our present study used the n-back tasks to determine if VOR responses could be reasonable measures for quantifying a person's MWLs.

## 2 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. This study integrated two parts of VOR models to estimate the VOR responses. The model, which is shown in Fig.1, has been proposed by Merfeld and Zupan [4]. The mechanism of a VOR model consists of three stages, as shown in Fig.1; (1) physical world & sensors, (2) signal processing of neural networks, and (3) generating command signal in angular velocity of eye movements. The VOR mechanism senses head movement; linear accelerations in three dimensions and angular velocities in three dimensions. Since this signal processing is not the same across people, identification of the model is required for an accurate simulation of each person. Eye movements are on three axes. However, this study utilized only the two main rotations: horizontal and vertical. Our study examined the discrepancies between the predicted eye movements by the model and the observed ones on these two axes.

The transformation from VOR output of the model to the angle is shown in Fig.2. The signal is modified through velocity storage, final common path and extraocular muscle [5]. Consequently, we integrate the model of the final common path and extraocular muscle with the VOR model described in Fig.1, and use it as the eye movement model of a subject.

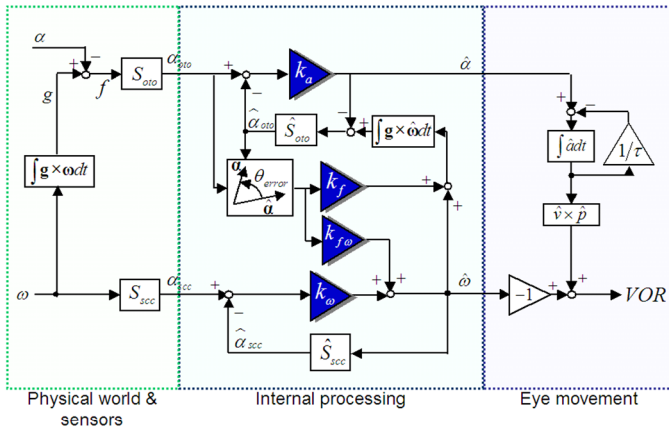


Fig. 1. Model of vestibulo-ocular reflex (VOR)

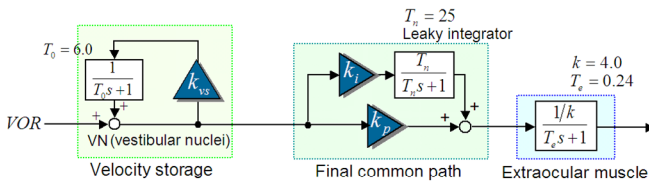


Fig. 2. Model of final common pass and muscle dynamics

### 3 Experimental Setup and Model Identification

The experiment set-up, sensors setting on the subjects, and definition of the coordinate system are shown in Fig.3. The rotating chair, which allowed rolling and pitching motions, was set up at a distance of 2[m] from the screen. The rotating chair, Joy Chair, provides rolling and pitching motions on the subject body. The subjects' head

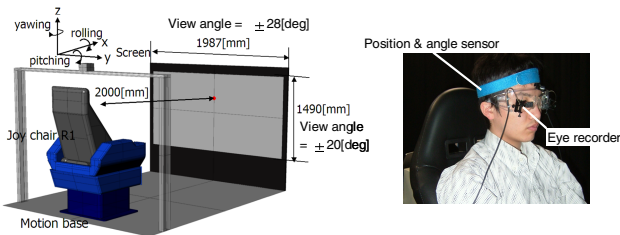


Fig. 3. Experimental setup and sensors setting

position and rotation angle, horizontal and vertical rotation angle of eyeball, rotation angle of the chair were measured at a sampling frequency of 50[Hz]. In the experiment for identifying personal dynamics of the model, the subjects were asked to stabilize their eyes on the center of screen so that the eye movements of VOR were induced by shaking the subject with moving chair. A person's VOR responses are almost perfectly predicted when the person is not engaging in a cognitively demanding task. The VOR model predicted the person's responses as perfectly as 0.92 in the coherence between the predicted VOR and the observed one.

We have applied genetic algorithm together with local search of gradient method to estimate the parameters in the VOR model which minimize the error expressed by

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

where  $\theta_{obs}$  is time series with three dimension sampled from the observed data of eye movement,  $\theta_{mdl}$  is time series estimated by the model. Model identification is the process to identify individual's eye movement dynamics that are represented by the 12 gain parameters shown with the seven triangle shapes in Fig.1 and Fig.2. These parameters are different by person and possibly by time. Before the identification, the measured eye movement has been treated with outlier removal process that extracts the rapid eye velocity data, which includes blinking and saccade. The identified model represents the VOR dynamics of a particular subject who was asked to put his gaze on the center of screen and was not engaging in any other demanding task.

One result of the identification is shown in Fig.4. The time responses of measured eye movements and the estimates from the model are compared in the figure for two cases. The left figure (I) shows the case when we applied the values for gain parameters as the initial values which have been suggested in Merfeld and Zupan [4] and Robinson [5]. It is shown with the right figure (II) that the better fitting of the predicted angular velocities to the measured ones are obtained after the identification.

We can use the frequency responses of the measured eye movements and the predicted ones as the index of concent between these signals. Specially, we take the coherence between these two signals for the index. The frequency responses shown in Fig.5 are the power spectra and the coherence between the measured eye movements

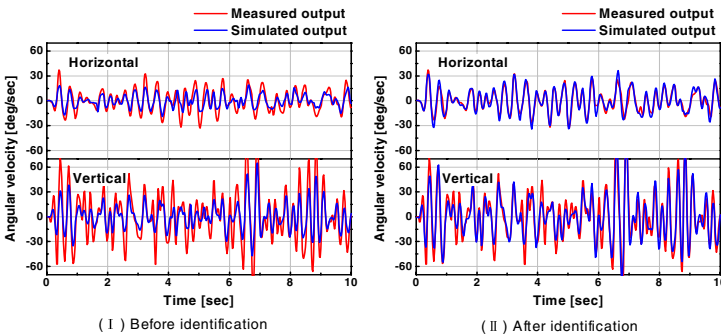
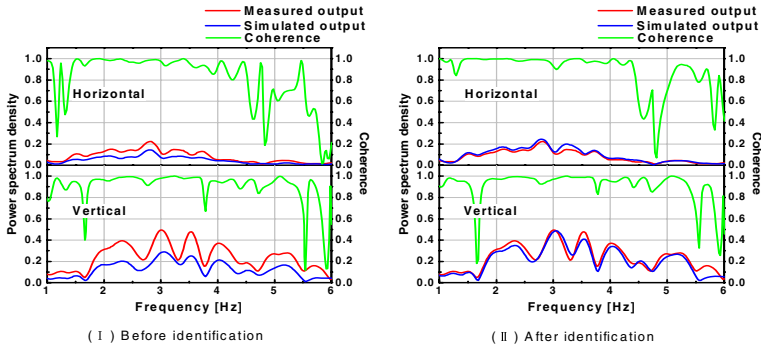


Fig. 4. Example of identification results (time responses)



**Fig. 5.** Example of identification results (frequency responses)

and the predicted ones from the identified models. In the figure, (I) and (II) correspond to the cases of before and after the identification. These results mean that the identified model is related linearly to the measured eye movements in the frequency range from 2 Hz to 4.5 Hz.

We conducted experiments with three subjects for identifying the models. The results show that the averages of predicted errors in three subjects were from 3 degree to 6 degree during eye movements of the range of more than 20 degree. We observed that the identified parameters vary in a certain range in three subjects. These identification results prove that the identified models can predict well the eye movements of the subjects.

## 4 Experimental 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 vestibulo-ocular reflex (VOR) responses from the identified model 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 Hz. The participants were asked to gaze upon a certain fixed point on a projector screen for 30 seconds on each trial. Then, VOR was continuously evoked to stabilize his point of gaze against disturbance on the head position during the task. Each of the 4 conditions was repeated three times for each participant. On each trial, eye movements and reaction time to every verbal presentation were recorded. The 4 conditions were as follows: 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, 1-back task, 2-back task, and 3-back task.

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

for each verbally presented letter when the presented letter matched the letter presented  $n$  items back in sequence, 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. The  $n$ -back tasks require to the subjects for holding /updating information, and decisions based on it. Working memory acts dynamically for such functioning during  $n$ -back tasks. Several researches of imaging brain functions by MRI have provided the observations that frontal association area, temporal association area, and Broca’s area activate during  $n$ -back tasks [6], [7]. 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.

## 5 Results

Overall, the results show that the more demanding MWL tasks induced more discrepancies between the predicted VOR responses with the identified model 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 participants took a longer time to answer in the more demanding  $n$ -back tasks, 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.

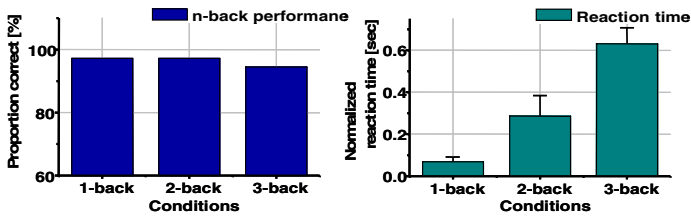


Fig. 6. Proportion correct in percentage and reaction time in the  $n$ -back tasks

Fig. 7 shows the results of the power spectrum (red and blue lines) and the coherence (green lines) between the measured VOR and the predicted VOR in the two conditions: the control condition (SRT) and 3-back conditions. The values of coherence are indicated in the green lines. However, the frequency responses above 4 Hz in the horizontal eye movements seemed to have too much noise and be at the error levels, not reflecting VOR responses much. The comparisons between the measured VOR and the simulated VOR in the four conditions in Fig.8 indicate several interesting things. First, there does not seem to be clear distinctions between the power spectrum curves in the vertical eye movements to distinguish the two conditions in Fig.7. The horizontal axis recorded similar power spectra. Second, conversely, the coherence curves (green lines) had relatively distinct characteristics from each other among the two conditions. The coherence curves of the control condition were higher on both

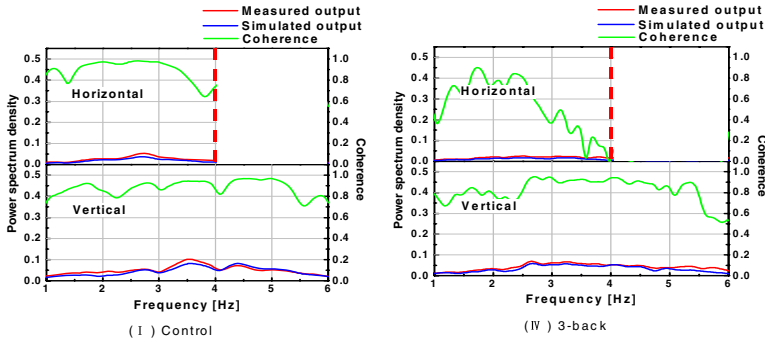


Fig. 7. Comparison of coherences and power spectrum in the tasks

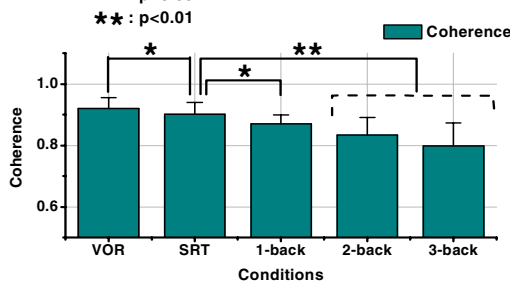


Fig. 8. Evaluating influence of working memory on coherence average (all subjects)

horizontal and vertical axes than the 3-back task. The task with the demanding mental workloads such as 3-back task had the lower coherence lines. Fig.8 shows the average of the coherence in the n-back tasks over all the three participants. This figure adds the VOR results, in which condition the participants were asked just similarly as in the identification experiment to stare at a fixed point on the projector screen without doing any additional task.

As expected, the identified model achieved the highest coherence under the condition of identification experiment, as high as 0.92, closest to the perfect coherence 1.00 among the 5 conditions. 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 [2], [8], [9]. These results suggest that VOR measures enable quantification of a MWL.

## 6 Extension to Optokinetic Situation

The model based method for quantitative evaluation of MWL can be extended to the cases when people use eye movements to pursue an object. The model of eye movements for such optokinetic cases can be described with the block diagram shown in Fig.9. The parts of VOR and final common path are the same as the sub-models described in chapter 2. We can use a mathematical model for the OKR part, which has been proposed by Schweugart, et al [10]. The model contains two components. They represent the dynamics of compensating functions for the retinal velocity error. One is a direct pathway characterized by a rapid rise in eye velocity, and the other is an indirect pathway that produces the slow-onset response. The error between the gaze direction and the position of the forward visual scene are inputs to the OKR model, and the model outputs the angular velocity of eyeball. The outputs are added to the outputs of VOR block; finally, the summation becomes the inputs of final common path.

The procedure for identifying the model of a particular subject is similar as VOR case. Only the difference is that there exist more parameters to be identified in the block of OKR. Thus, we can obtain the model which represents the dynamics of a person’s eye-movements in extensive situations. If the model is identified for the case that the subject works without any heavy MWL, the predicted output from the model can be used to compare with the real eye movements of the subject.

The same experimental setup which is shown in Fig.1 can be used for optkinetic situations. We add a steering on the chair, and make it work as a driving simulator. An example of experimental display of optkinetic situation is shown in Fig.10. The

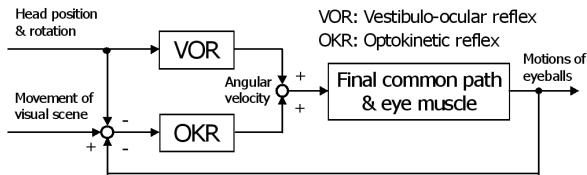


Fig. 9. Block diagram of reflex eye-movement model (general case)

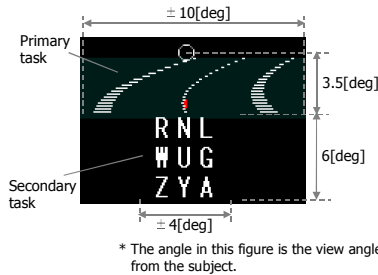
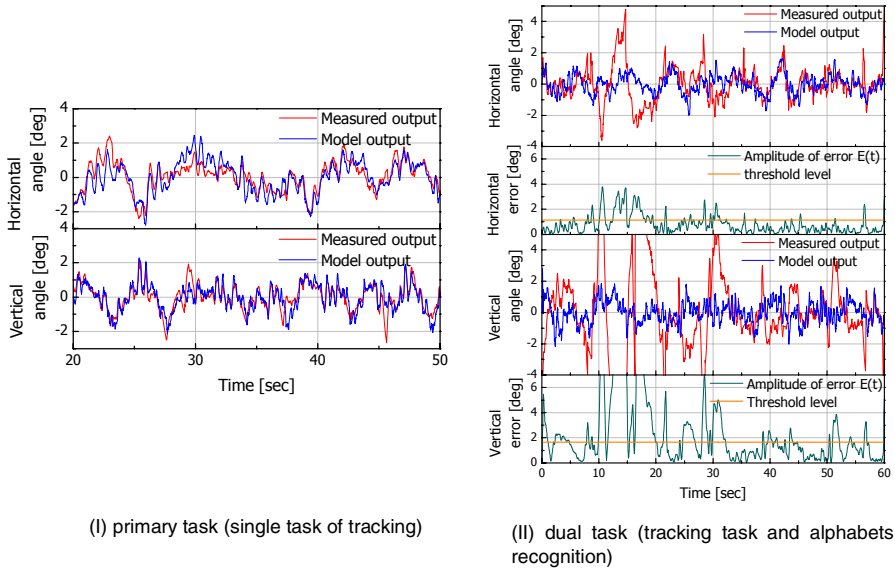


Fig. 10. Example of experimental display





**Fig. 11.** Comparison of measured angles with predicted angles by the identified model

variation between target signal (white line in the center) and controlled signal (red marker) is displayed on the screen. The subjects were required to manipulate the steering to bring the center line of the displayed road close to the center of the screen. The target signal was colored noise that is made from gaussian white noise. It has frequency band between 0.03[Hz] and 0.5[Hz]. In addition, a circle displayed on the center of screen randomly changes to red or blue or yellow. The subjects was required to detect the change of color and push the button that is placed on the left side of the steering as quickly as possible when the circle changes to the red.

To examine the influence of MWL, we asked the subjects to conduct a dual task: a primary tracking task with alphabets recognition as the secondary task. The secondary task was for the subjects to detect particular alphabets from the displayed alphabets. The displayed alphabets include one particular alphabet that subjects were asked to detect. Alphabets were randomly displayed on the screen for one second once every 3[sec] to 6[sec] as shown in Fig.10. The predicted eye movements by the identified model for the primary task are compared with the measured eye movements in Fig.11 (I). It shows a good content between the measured eye movements and the predicted ones. On the other hand, in the case of the dual task the predicted eye movements which were calculated from the identified model for the single task (primary tracking task) have considerable errors to the measured eye movements, which is shown in Fig.11 (II). This result suggests that we can quantify the effect of MWL by taking the norm of the absolute error for this kind of optokinetic situation. Another research in our laboratory examines the real-time aspect of this model to answer how accurately the person's MWL can be quantified using model-based eye-movement measures in real-time [11]. Possible applications of this eye-movement method are many since this method can objectively quantify a person's MWL in real-time without using large devices.

## 7 Concluding Remarks

It is shown in this study that VOR measures can be used as a method to objectively quantify a person's MWL. The results showed 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. Additionally, the extension to optokinetic cases has been given, which shows the possibility of this model-based method for quantifying MWL in practical situations.

## References

1. Stanton, N.A., Salmon, P.M., Walker, G.H., Baber, C., Jenkins, D.P.: *Human Factors Methods – A Practical Guide for Engineering and Design*. Ashgate, Burlington (2005)
2. Furman, J.M., Müller, M.L.T.M., Redfern, M.S.: Visual-vestibular stimulation interferes with information processing in young and older humans. *Experimental Brain Research* 152, 383–392 (2003)
3. Shibata, N., Obinata, G., Kodera, H., Hamada, H.: Evaluating the influence of distractions to drivers based on eye movement model. In: *FISTA World Automotive Congress 2006*, F2006D164 (2006)
4. Merfeld, D.M., Zupan, L.H.: Neural processing of gravito-inertial cues in humans. III. Modeling tilt and translation responses. *Journal of neurophysiology* 87, 819–833 (2002)
5. Robinson, D.A.: The use of control system analysis in the neurophysiology of eye movements. *Annual review of neuroscience* 4, 463–503 (1981)
6. Braver, T.S., et al.: A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage* 5, 49–62 (1997)
7. Cohen, J.D., et al.: Activation of prefrontal cortex in a non-spatial working memory task with functional MRI. *Human Brain Mapping* 1, 293–304 (1994)
8. Yardley, L., Gardner, M., Lavie, N., Gresty, M.: Attention demand of perception of passive self-motion in darkness. *Neuropsychologia* 37, 1293–1301 (1999)
9. Talkowski, M.E., Redfern, M.S., Jennings, J.R., Furman, J.M.: Cognitive requirements for vestibular and ocular motor processing in healthy adults and patients with unilateral vestibular lesions. *Journal of Cognitive Neuroscience* 17(9), 1432–1441 (2005)
10. Schweigart, G., et al.: Eye movements during combined pursuit, optokinetic and vestibular stimulation in macaque monkey. *Experimental Brain Research* 127, 54–66 (1999)
11. Obinata, G., Usui, T., Shibata, N.: On-line method for evaluating driver distraction of memory-decision workload based on dynamics of vestibulo-ocular reflex. *Review of Automotive Engineering* 29, 627–632 (2008)