PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

Effects of localization control of warning sound combined with visual information in vehicle cockpit

Kiichi NAKA[1]; Katsuya YAMAUCHI[2]; Nobuaki TANOUE[3]; Ayumu KAWATA[3]

[1] Graduate School of Design, Kyushu University, Japan

[2] Faculty of Design, Kyushu University, Japan

[3] Pioneer Corporation, Japan

ABSTRACT

Recently, many kinds of information are provided to drivers because of the improvement of ADAS (Advanced Driver Assistance System). Amount of the information in vehicle cockpit seems to be increased due to the development of automobile technology such as autonomous driving. It is very important to consider the design of appropriate auditory information to avoid overlooking visual information cues. In this paper, we examined the effect of localization control of warning sounds to determine the direction of visual information using a driving simulator. The number of overlooked visual information cues and subjective task load were significantly decreased by the localization control in comparison with the no-localization and no-sound conditions. A driver could recognize the direction of visual information without removing their gaze from frontal view so that a driver also could find the visual information rapidly and accurately, and could watch the frontal view more long duration. It led to steady driving performance such as steering and speed control. In contrast, these effects were diminished if the timbre of the signal was difficult to localize.

Keywords: Warning sound, Driving simulator, Subjective evaluation, Gaze, Driving behavior

1. INTRODUCTION

With recent developments of ADAS (Advanced Driver Assistance System), vehicles has many kinds of information and the amount of information provided to drivers are increasing. Also, the methods to provide information are diversifying. The information will be delivered on not only the instrument panel but also various positions. The door mirror indicator light (1) is one of the example that already exist in the car commercially available.

The increasing and diversifying of information in vehicle cockpit have possibility to make the drivers confused. It could disturb the primary driving operation. Therefore, HMI (Human Machine Interface) to present information has to be appropriately designed. Jakus et al. (2) showed that, in a case of the drivers have secondly task other than driving, they can perform the tasks more rapidly and accurately by providing multimodal feedback that integrates audio-visual information than providing each visual or auditory solely. Then, this paper focused on audio-visual information to the drivers in vehicle cockpit.

The audio information has an advantage to be able to recognize the provided direction regardless of people's gazing direction. The drivers have to gaze the front while driving. It is thought that audio information such as warning sound can tell the direction of visual information at the various positions in vehicle cockpit by using localization control. It can be expected that the driver can find visual information rapidly while driving by telling the direction the driver has to gaze.

This paper examined comprehensively the effects of localization control of warning sounds telling the direction of visual information through a laboratory experiment using driving simulator through the evaluation of accuracy of finding visual information, gaze, driving performance, and subjective mental workload.

[1] kn.pos222@icloud.com
[2] yamauchi@design.kyushu-u.ac.jp

2. EXPERIMENTAL SYSTEM

Figure 1 shows the overview and the diagram of the experimental set-up. The system consisted of three units; driving simulator unit, visual information unit, and warning sound unit. In addition, the participants' faces were recorded using a video camera from their front.

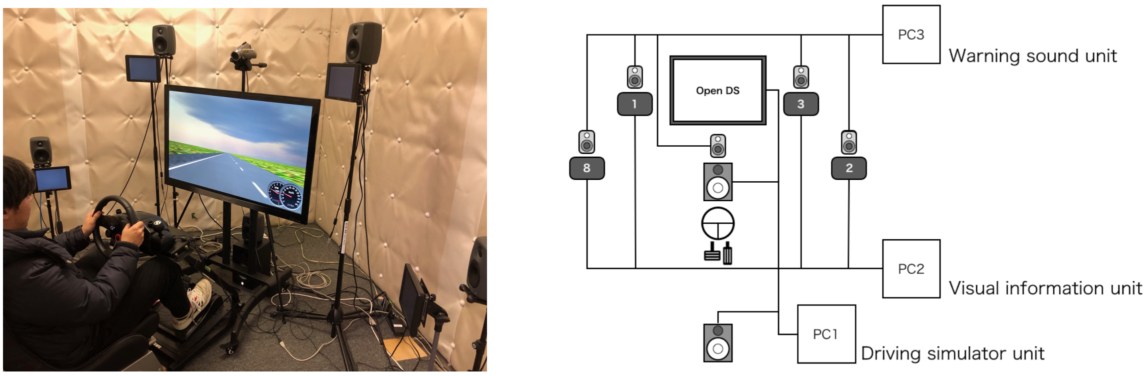


Figure 1 – Overview (left one) and diagram (right one) of the experimental set-up.

2.1 Driving simulator unit

"OpenDS" (3), which is JAVA-based open-source driving simulator, was used in the experiment. It has multiple scenarios, and "MotorwayTask" scenario was chosen for the experiment. The screenshot of "MotorwayTask" is shown in Figure 2. The scenario consists of two-lanes straight road and no objects nor the other cars.

The participants manipulated steering wheel and speed control device (Logitec/ Driving Force GT). A vehicle's speed [km/h] and engine rotation [rpm] were displayed on the main monitor which is at the participant's frontal. The distance between the center of the main monitor in front of the participants. The distance between the center of the main monitor and steering wheel was about 1.2 m. The sound imitated the vehicle cockpit noise was given from loudspeakers which located below the main monitor and behind the driver seat. The equivalent continuous A-weighted sound pressure level over the period of 10 s ($L_{\mathrm{Aeq,10s}}$) were 56.2 dB while idling and 56.2 dB while driving at constant speed of 60 km/h. The vehicle's coordinate position and speed in driving simulator were recorded as logging data every 50 ms.

Figure 2 – Screenshot of "MororwayTask" scenario.

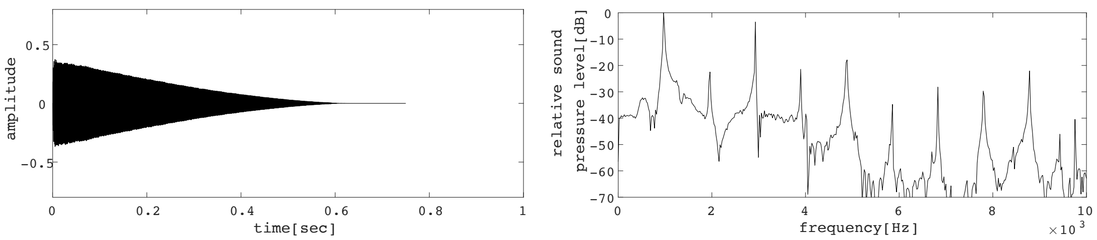
2.2 Visual information unit

As the abstraction of visual information that the drivers have to find while driving, the single digits (0–9) were presented in four small monitors around the main monitor. The small monitors were placed at -50˚, -30˚, +30˚, and +50˚ where the viewing angle on the horizontal plane of the main monitor from the driver's seat was 0˚. The distance between the center of small monitors and steering wheel were set as 1.2 m. The height from the floor was 1.3 m for the monitors at -30˚ and +30˚, and 0.65 m for those at -50˚ and +50˚.

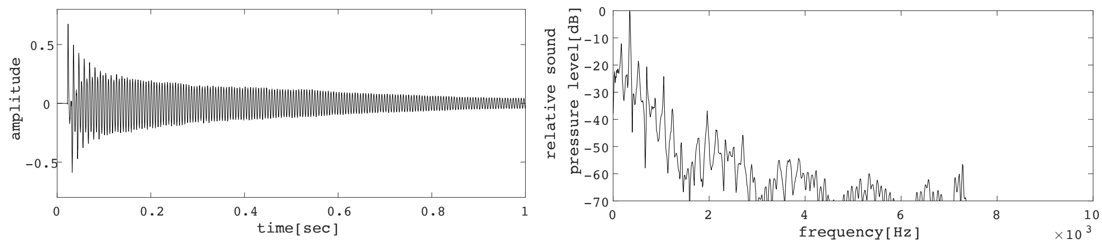
The single digits in white were changed randomly each small monitor in 1 s cycle, irregularly one of the digit was shown in red as to assume some high priority information. The height of the digit on the monitor was 18 mm.

2.3 Warning sound unit

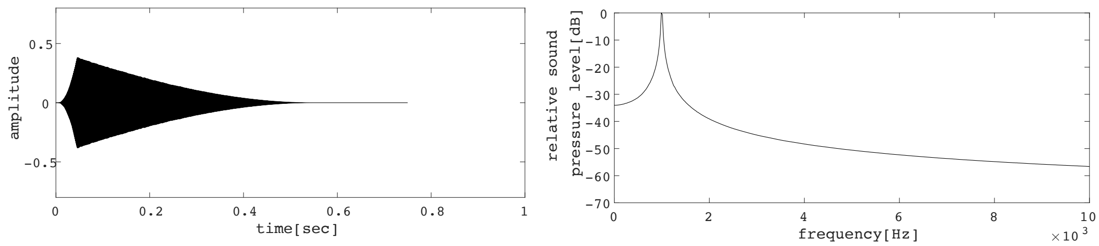
Warning sound was presented from the loudspeaker located aside of each monitor when the red digit appeared. Three kinds of timbre (A, B, and C) were prepared as the warning sounds. The waveform and frequency spectrum are shown in Figure 3. The fundamental frequency of the warning sound A was about 1 kHz and it contained multiple integral harmonics. The warning sound B was collected from a vehicle navigation system commercially available, and its fundamental frequency was about 200 Hz and contained multiple integral harmonics. The warning sound C consisted of only one frequency component at 1 kHz. The warning sound A and C had similar amplitude envelope of a sharp rise and a decay of approximately 0.5 s. The warning sound B had a sharp rise and longer decay of approximately 1 s. Only the warning sound A was used in experiment 1, and the warning sound B and C were used in experiment 2. A-weighted maximum sound level (L_{Amax}) at the driver's head position was adjusted to be 66.2 dB for all warning sounds.



(a) Warning sound A



(b) Warning sound B



(c) Warning sound C

Figure 3 – Waveform and frequency spectrum of each warning sound.

3. EXPERIMENT 1

Experiment 1 examined the effects of localization control of warning sounds on the accuracy of finding visual information, gaze, driving behavior, and subjective mental workload.

3.1 Tasks

The participants were asked to take “finding digit task” and “driving task”.

In “finding digit task”, the participants were asked to find the red digit on four small monitors and answer it verbally as quickly as possible. The duration of one trial was 90 s, and the red digit appeared 20 times in one trial.

In “driving task”, the participants were asked to manipulate the driving simulator with obeying two rules; to keep the driving speed in 60 km/h, and to stay at the center of left side lane.

3.2 Procedure

Experimental conditions are shown in Table 1. Three conditions were set in a difference of how to provide warning sound when the red digit appeared. The condition 1 provided no warning sound, the condition 2 provided warning sounds from a single loudspeaker below the main monitor, and the condition 3 provided warning sounds using localization control from the loudspeakers which were at located aside of each small monitor. The participants were asked to take “finding digit task” and “driving task” simultaneously in these conditions. As a control, the participants were asked to take only “driving task” in condition 0.

Table 1. Conditions of experiment 1.

	task	warning sound	loudspeaker position
condition 0	driving	–	–
condition 1	finding digit and driving	no sound	–
condition 2	finding digit and driving	warning sound A	below the main monitor
condition 3	finding digit and driving	warning sound A	aside of each small monitor

Before starting trials, the participants practiced manipulating driving simulator at least 5 minutes to became accustomed to the experimental environment. Two practice trials and two actual trials were conducted for each condition. The sequence of conducting conditions was randomized in each participant. After finished each condition, participants were asked to evaluate their subjective mental workload on the scale.

3.3 Evaluation

The error rate, which was the ratio of missed red digit to all 20 red digits in one trial, was calculated to evaluate the accuracy of finding visual information.

The subjective mental workload was evaluated on four scales based on DALI (4) (Driving Activity Load Index); the effort of attention, visual demand, auditory demand, and interference. The DALI was a subjective mental workload evaluation method to evaluate the subjective burden when driving a vehicle based on NASA-TLX (5) (NASA Task Load Index). The evaluation obtained 0 to 100 by marking an arbitrary point on the line segment with “low” at the left end of the line and “high” at the right end, and the distance from the left end of the line to the evaluation position was standardized by the line length.

In the evaluation of gaze, the moving picture of one trial from the front of the participants was sampled every 10 ms, and the direction of gaze was classified into six types; the front main monitor, four small monitors, and unknow. The average gaze reach time T_r [s] taken to reach the direction of the red digit after the appearance of the red digit was evaluated. The duration spent looking at the front during one trial was evaluated as the duration of front center gaze time T_c [s]. With regard to the average gaze reach time T_r [s], the cases that the gaze did not reach to the direction of the red digit within 1 s and the gaze was at the direction of the red digit from the moment they appeared were excluded. In order to examine the difference in auditory information added conditions, the gaze was evaluated for condition 2 and 3.

The steering wheel and speed fluctuation were evaluated from the coordinate position of the vehicle body from the travel log and the travel speed. The steering fluctuation U_{steer} [m] was determined by equation (1) using the unbiased standard deviation of the x-coordinate position representing the road crossing direction.

$$U_{steer} = \sqrt{\sum_{i=1}^{n} \frac{(x_i - x_m)^2}{n-1}} \tag{1}$$

where x_i was the x-coordinate of the vehicle center at a certain point i, and x_m was the average value of x-coordinates during one trial. Similarly, the speed fluctuation U_v [km/h] was determined by equation (2) using the unbiased standard deviation of the speed v. The v_i was the traveling speed at a certain point i. The v_m was the average value of the traveling speed during one trial.

$$U_v = \sqrt{\sum_{i=1}^{n} \frac{(v_i - v_m)^2}{n-1}} \tag{2}$$

3.4 Participants and accuracy of their sound source localization

Ten males and 10 females participated in this experiment. The participants ranged in the age from 21 to 25 years old. All the participants had the normal hearing ability in self-report. They have had a driver licenses for the period ranged from 7 months to 6 years.

To investigate the participants could accurately realize the direction of the sound source, a preliminary experiment was conducted before the experiment. The participants asked to answer which loudspeaker played warning sound. The warning sounds were presented 40 times for each participant. The average of correct answer rate was 99.5% (S.D.=1.0). Then, it was confirmed that the all participants could identify the direction of warning sounds.

4. RESULTS AND ANALYSIS (EXP. 1)

4.1 The accuracy of finding visual information

The error rate of each condition is shown in Figure 4. The error rate of three conditions were compared using non-parametric multiple comparison with Bonferroni correction ($\alpha = 0.05$). The analysis showed significant differences between all conditions.

The visual information was overlooked more than 50% when using no warning sound (condition 1). The number of overlooked visual information was diminished using the warning sound. However, more than 20% of visual information was still overlooked when using a single loudspeaker (condition 2). By telling the direction of visual information using localization control (condition 3), visual information became to be found accurately.

4.2 Subjective mental workload

The subjective mental workload of each condition is shown in Figure 5. The subjective mental workload of four conditions were compared using non-parametric multiple comparison with Bonferroni correction ($\alpha = 0.05$). In the effort of attention, the analysis showed significant differences between all conditions except between the condition 1 and 2. In the visual demand and interference, the analysis showed significant differences between all conditions. In the auditory demand, the analysis showed significant differences between the condition 0 and 2, 0 and 3, 1 and 2, and 1 and 3.

The effort of attention, visual demand, and interference were diminished using localization control than using a single loudspeaker. It showed that localization control of warning sound could reduce the demand of finding visual information. The auditory demand was higher when using a single loudspeaker than using no warning sound, while the auditory demand was lower when using localization control than using a single loudspeaker. According to Schwarz et al. (6), specific information had some positive effects to driver. It seemed that auditory demand was diminished by using more helpful warning sounds which indicated the direction of visual information. The number of channels of loudspeaker was seemed to be irrelevant to the auditory demand.

4.3 Gaze and driving behavior

The average gaze reach time T_r and the duration of front center gaze time T_c are shown in Figure 6. Thirty-three data were obtained under condition 2, and 40 data were obtained under condition 3 as T_r. The T_r of two conditions were compared using t-test. The T_r under condition 3 was significantly shorter than condition 2 ($t(19) = 19.49, p < 0.05$). The T_c of two conditions were compared using t test. The T_c under condition 3 was significantly longer than condition 2 ($t(19) = -5.98, p < 0.05$).

The steering fluctuation U_{steer} and speed fluctuation U_v of each condition are shown in Figure 7. The U_{steer} and U_v of four conditions were compared using non-parametric multiple comparison with Bonferroni correction ($\alpha = 0.05$). The analysis showed significant differences between all conditions except condition 0 and 3 in the U_{steer}, and between condition 1 and 0, and 2 and 0 in the

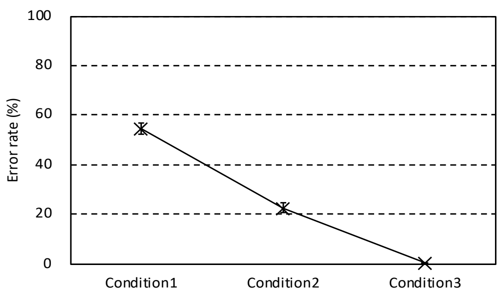


Figure 4 – Error rates on “finding digit task”.

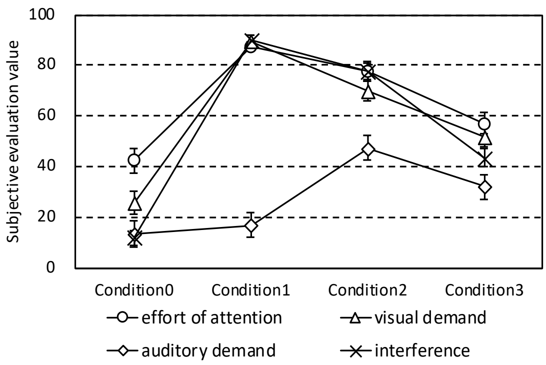


Figure 5 – Subjective mental workload.

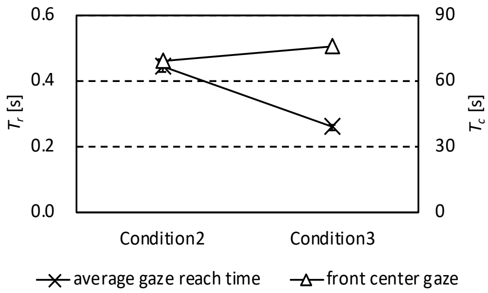


Figure 6 – Average gaze reach time T_r and front center gaze T_c.

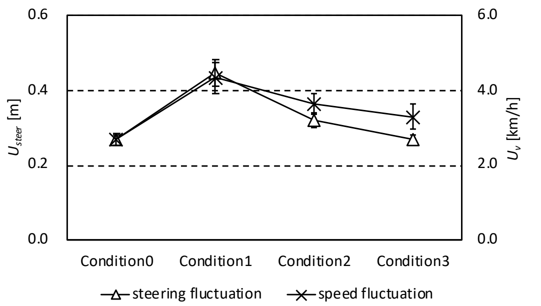


Figure 7 – Steering fluctuation U_{steer} and speed fluctuation U_v.

U_v.

By giving the direction of visual information using localization control, average gaze reach time T_r was shortened. The driver could recognize the direction of the visual information while watching the front by the localization control. Therefore, it is considered that useless eye movement was reduced and visual information can be reached at the shortest distance. The reduction of useless eye movement also led to an increase in time to look at the front.

The steering fluctuation of using localization control condition was similar to performing only driving (condition 0). It was thought that the increase of frontal gaze led to stable driving. The duration of frontal gaze was decreased when using a single loudspeaker and it led to increasing steering fluctuation. With regard to the speed fluctuation, it tended to be more stabilized by localization control.

4.4 Conclusion of exp. 1

It is found that the driver could find visual information accurately and perform more stable driving operation with low burden by telling the direction of visual information using localization control of warning sound combined with visual information.

5. EXPERIMENT 2

In experiment 2, the effects on difference of the timbre was verified when the direction of visual information was given using localization control of warning sound. Tasks and evaluation were similar to experiment 1.

5.1 Procedure

Experimental conditions are shown in Table 2. Two conditions were set in difference of the timbre of warning sound. The warning sounds were provided from the direction of the red number using localization control in both conditions. Similar to experiment 1, the participants practiced manipulating driving simulator to became accustomed to experimental environment before starting trials. Two practice and two actual trials were conducted in each condition. The sequence of conducting conditions was randomized in each participant. After finished each condition, participants were asked to evaluate their subjective mental workload on the scale.

Table 2. Conditions of experiment 2.

	task	warning sound	loudspeaker position
condition 4	finding number and driving	warning sound B	aside of each small monitor
condition 5	finding number and driving	warning sound C	aside of each small monitor

5.2 Participants and accuracy of their sound source localization

Five males and 4 females participated in this experiment. The participants ranged in the age from 22 to 25 years old. All the participants had the normal hearing ability in self-report. They have had a driver licenses in the period ranged from 1 year to 5 years and 1 month.

To investigate the participants could realize accurately the direction of the sound, a preliminary experiment was conducted before the experiment. The participants asked to answer which loudspeaker played warning sound. Each warning sound were presented 40 times. The average of correct answer rate of warning sound B was 98.1% (S.D.=2.4) and warning sound C was 58.9% (S.D.=16.5). It was confirmed that warning sound B was identified its direction by the participants and the warning sound C was difficult to be localized.

6. RESULTS AND ANALYSIS (EXP. 2)

The error rate of the warning sound B and C are shown in Figure 8. The error rate of warning sound A of experiment 1 is also shown in the figure. The error rate of three warning sounds were compared using non-parametric multiple comparison with Bonferroni correction ($\alpha = 0.05$). The analysis showed significant differences between the warning sound A and C, and B and C.

The subjective mental workloads of each condition are shown in Figure 9. The subjective mental workload of three warning sounds were compared using non-parametric multiple comparison with Bonferroni correction ($\alpha = 0.05$). In the effort of attention, the analysis showed significant differences between the warning sound A and C, and B and C. In the visual demand and the interference, the analysis showed significant differences between the warning sound A and C. In the auditory demand, there were no significant differences.

The steering fluctuation U_{steer} and speed fluctuation U_v are shown in Figure 10. The condition 0 of experiment 1 is also shown in the figure. The steering fluctuation U_{steer} and speed fluctuation U_v of each condition were compared using non-parametric multiple comparison with Bonferroni correction ($\alpha = 0.05$). There were no significant differences in each warning sound.

The significantly difference were not found in driving behavior. However, the warning sound C which was difficult to be localized could not assist the driver to find the visual information sufficiently even if using localization control. Also, subjective mental workload such as the effort of attention, visual demand, and interference were significantly higher than the others. It was found that if the timbre was difficult to be localized, warning sound could not help the driver sufficiently to find the visual information because the driver could not recognize the direction by given warning sound.

7. CONCLUSIONS

The effects of telling the direction of visual information to the driver using localization control of warning sound were examined using driving simulator.

The result of the first experiment showed that, by telling the direction of visual information using localization control of warning sound, the number of overlooked visual information was reduced and the steering and speed control became stable. The average gaze reach time to visual information was shortened and the duration of driver's frontal gaze was increased. Also, the driver's subjective mental workload was diminished.

Warning sound could tell the direction of visual information by localization control in addition to the timing of visual information appeared. Therefore, the drivers could understand "when" and "where" to look while watching the front. Even in the case which visual information was not in their gaze, the drivers could gaze the visual information in a short time without looking unnecessary direction. The reduction of useless eye movement during driving led to an increase of frontal gaze duration. It enables the drivers to concentrate on driving operation. Therefore, the control of steering wheel and speed became stable while confirming visual information. Also, the subjective mental workload could be reduced.

The effects on timbre was examined in the experiment 2 for the case of using localization control. The number of overlooked visual information and the driver's subjective mental workload were increased when the timbre was difficult to be localized. If the timbre was difficult to be localized, warning sound could not

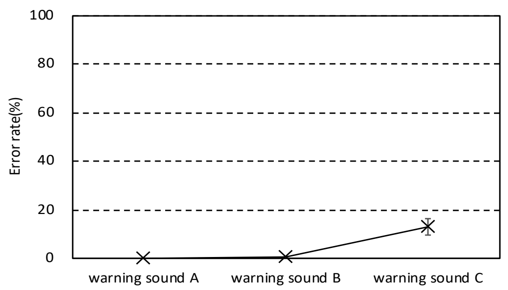


Figure 8 – Error rates on “finding digit task”.

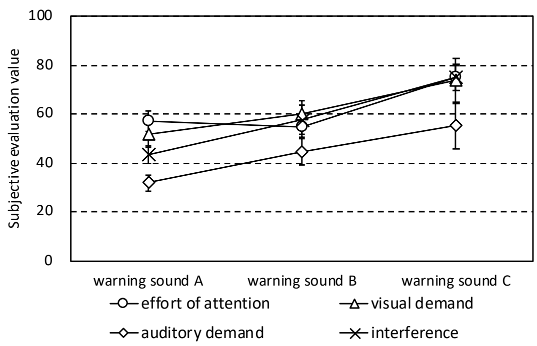


Figure 9 – Subjective mental workload.

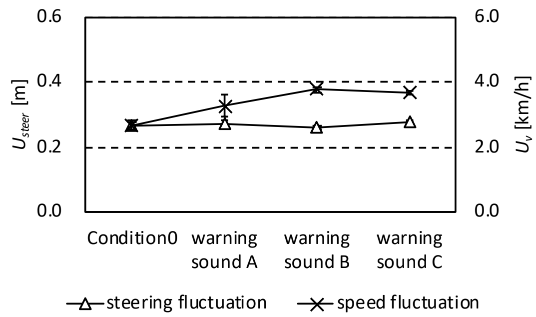


Figure 10 – Steering fluctuation U_{steer} and speed fluctuation U_v.

tell the direction even if using localization control. It led increase of missing visual information and subjective burden to driver. It was necessary to use the timbre that the drivers could recognize the provided direction when using localization control as an HMI.

REFERENCES

1. T. Kanamaru, H. Kuwamoto. The transition of HMI technology and the approach for in-vehicle infotainment system. Journal of Society of Automotive Engineers of Japan. 2015; 69 (3); pp. 39–42.
2. G. Jakus, C. Dicke, J. Sodnik. A user study of auditory, head-up and multi-modal displays in vehicles. Applied Ergonomics. 2015; 46; pp. 184–192.
3. “openDS,” https://opends.dfki.de. [Accessed: 31-May-2019].
4. A. Pauzie. A method to assess the driver mental workload: the Driving Activity Load Index (DALI). IET Intelligent Transport System, 2008; 2; pp. 315–322.
5. S. G. Hart, L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, Advances in Psychology, 1988; 52; pp. 139–183.
6. F. Schwarz, W. Fastenmeier. Augmented reality warnings in vehicles: Effects of modality and specificity on eggectiveness. Accident Analysis and Prevention; 2017 101; pp. 55–66.