Human Interaction and Emerging Technologies (IHIET 2023), Vol. 111, 2023, 518–525
https://doi.org/10.54941/ahfe1004058

The Effect of Auditory Stimulus on Gazing Behavior in Driving Assistance Systems

Kisa Takao, Hironori Uchida, Yujle Li, and Yoshihisa Nakatoh

Kyushu Institute of Technology, 1-1, Sensui-cho, Tobata-ku, Kitakyushu-shi, Fukuoka, Japan

ABSTRACT

In order to prevent traffic accidents caused by drivers' inattention, various driving assistance systems have been developed. These systems alert the drivers through sound and warning displays. However, since the type of sound and the location of the sound differs depending on the car model, if the driver does not pay attention, may miss the warning signs. This study examines a method to support drivers' perception of danger by presenting auditory stimuli to promote visual awareness of an arbitrary location. A previous study showed that when a simple visual task was used, the presentation of stimulus sound accelerated the selection of the next fixation target. There were differences in confirmatory behavior toward the target depending on the direction of sound production. Based on these results, we hypothesized that when watching actual driving images, the differences in gaze behavior would emerge depending on whether sounds were presented from the target direction and the type of sound and measured subjects' gaze. The results obtained indicate that it is possible that auditory stimuli can induce gaze. However, the effect tends to be stronger when an eye-catching object appears in front of the eyes, and it is not clear whether stimulus sound alone is sufficient to control the eye's gaze in detail. And, it is difficult to compare the differences in induction effects among the four different stimulus sounds since the results also suggest that visual information is preferred for attention in this study. In a subjective evaluation conducted after the experiment, more than half of the subjects responded that their attention was directed toward the direction of the sound presentation. Based on the responses of the subjects to the questionnaire, we would like to investigate the future relationship between sound representation methods and features of visual stimuli and induction effects.

Keywords: Driver, Saccadic eye movement, Stimulus sound

INTRODUCTION

The number of casualties in traffic accidents has decreased by the growth of Advanced Driver-Assistance Systems (ADAS). However, the total number of traffic accidents remains high, and moreover, it is reported that accidents caused by drivers' inattention, such as careless driving, account for approximately half of the total number, according to Cabinet Office (Cabinet Office).

Providing audible warnings of detected hazards based on information acquired by sensors is a common approach in driver assistance systems. Another approach is to visually present the danger by displaying warning lights or warning text on the meter panel. However, the warning methods of these support systems differ by vehicle type in terms of warning display position, the type of warning sound, and the location of the warning sound generation. In addition, while warning displays have the advantage of easily communicating the content of warnings to drivers, there is the problem that they are difficult to notice when drivers themselves are not paying attention to their surroundings, such as when they are driving while looking aside, and there is the possibility that they may miss the warning. Therefore, this study examines an approach that supports the perception of danger by guiding the driver's gaze through the representation of sound.

Eye movements and human attention are thought to be closely related, and there has been considerable research on these relationships. Shifts in the line of sight are caused by rotational movements of the eyeball caused by the external eye muscles (eye movements). Saccadic eye movement is a type of eye movement that consists of two movements: an eye movement called a saccade and a gazing behavior called a fixation. During fixation, the observer is thought to be simultaneously observing the gaze target and determining the next saccade target. In particular, for the latter, it is inferred that people evaluate visual information and determine the next fixation position to be the one that they judge to have the highest priority for attention (John, M. F. and Iain, D. G. (2006)). Moreover, the effect of auditory stimuli on eye movements has been studied in fundamental research areas. In a simple visual search task, previous studies have confirmed that the presentation of stimulus sounds and auditory cues speeds up the reaction time for saccades to the gazing target (Douglas, P. Munoz, B. D. (1995)) and that there are differences in the confirmation behavior toward the gazing target depending on the direction of the sound presentation (Ito, K., Yamanaka, T., Goto, M. and Miki, Y. (2014)).

Therefore, we consider that it may be possible to direct the gaze to an arbitrary position to which we wish to direct attention by presenting sensory stimuli before the saccade localization position is determined during fixation. In this study, we will investigate whether auditory stimuli can promote visual awareness of the object to which we wish to direct attention.

METHOD

Referring to previous studies, we hypothesize that differences in gaze behavior will emerge depending on the presence or absence of sound presentation from the target direction and the type of sound, and conduct experiments. To measure eye behavior under conditions similar to the actual driving environment, subjects watch driving videos taken from the driver's seat. Eye line-of-sight measurements are taken twice for each running video: One condition with stimulus sound and another condition without stimulus sound. Stimulus sounds are presented in the direction of the target that requires attention. Eye movements are measured using Tobii Pro Glass 3 (sampling rate 100 Hz, Tobii).

In the experiment, four different stimulus sounds are used to compare the differences in gaze behavior with and without sound presentation. Then, we analyze based on saccadic eye movements, whether differences in gaze behavior are produced by the presence or absence of stimulus sounds, and whether there are differences in the effects of different types of stimulus sounds on gaze behavior.

Figure 1 shows the layout of the interior of the soundproof chamber where the experiments will be performed. All Subjects (10 men, aged 21-24 years) hold a driver's license. All subjects are tested to ensure that they have normal hearing and the necessary visual acuity for driving to ensure that there are no obstacles to experimenting. Also, if subjects use vision correction devices, limit them to contact lenses only, to reduce measurement error in the eye-tracking camera used.

Figure 1: Experiment environment.

The conditions used in the experiment are shown in Table 1. Four types of stimulus sounds will be used. They are white noise (8 kHz LPF) and synthetic waves (1.1 kHz square wave and 2.2 kHz sine wave), which have been shown in previous studies to have good localization properties (Morikawa, D. Shimakura and N. Hirahara, T. (2009)) (Ito, K., Yamanaka, T., Goto, M. and Miki, Y. (2014)); sine waves (2 kHz) and instrumental sound, glockenspiel (2.1 kHz), which are sounds with frequencies near 2 kHz and are thought to be less likely to cause hearing loss due to aging and masking (Association for Electric Home Appliances. (2018)). The stimulus sound is presented for 1 second, 700 ms before the time at which the gaze should begin, and the volume is set to 60 dB.

Table 1. A table of experimental conditions.

	Driving video 1	Driving video 2	Driving video 3	Driving video 4
auditory stimulus presentation conditions	no stimulus	no stimulus	no stimulus	no stimulus
(speaker number)	white noise (5)	sine wave (3)	synthetic waves (5)	glockenspiel (3)

Driving videos will be edited from those published by JAF on YouTube. One driving video contains six different dangerous scenes (JAF).

The eye gaze is measured when four different stimulus sounds are presented, but when the same content is viewed several times, the subject's gaze may be guided by the memory of the images. Therefore, the following three measures shall be taken.

1. Driving videos with completely different content are used for each stimulus sound. (Therefore, eye movement has measured a total of eight times.)
2. The same driving video is not presented two times straight.
3. Only one of the six danger scenes in the driving video is used for analysis: content related to "jump-out". This is not told to the subjects.

The data obtained from the measurements is used for comparison between stimulus presentation condition and no stimulus condition. The two comparison categories are saccade frequency and induction time. Saccade frequency is the number of saccades occurring per second. It is used to investigate whether the presentation of sound affects the occurrence of saccades. The induction time is the time between the presentation of the stimulus sound and the start of fixation on the target or area specified. If the time is shortened by the sound presentation, it is considered induced.

We define two types of induction times. The first is the induction time to the gazing area: the time when fixation begins on the peripheral area where the gazing target appears, and the second is the induction time to the gazing target: the time when fixation begins directly on the gazing target.

The above items are analyzed using gaze measurement data for 2000 ms after the stimulus tones are presented. Saccades and fixations that occur before the stimulus sounds are presented are not included in the analysis. For the analysis of saccadic eye movements, the analysis software Tobii Pro Lab is used. Saccades are determined when the movement speed is 100 deg/sec or more, and fixations are determined when the movement range is within 0.5 deg and persists for 60 ms or more.

EXPERIMENTAL RESULTS

A comparison of the difference in saccade frequencies among the four stimulus sounds is shown in Figure 2. Since the difference in saccade frequency averaged less than 0.5, auditory stimulation by the stimulus sound had a small effect on the occurrence of saccades.

Next, the results for induction time are shown in Figure 3 and Figure 4.

In the condition in which white noise and sine waves were presented, the induction time to the gaze region was significantly shortened and to the gaze target was not. And, there was no significant reduction in the induction time to the gaze area. In contrast, there was no significant reduction in the induction time to the gaze area when synthetic wave and instrumental sound (glockenspiel) were presented.

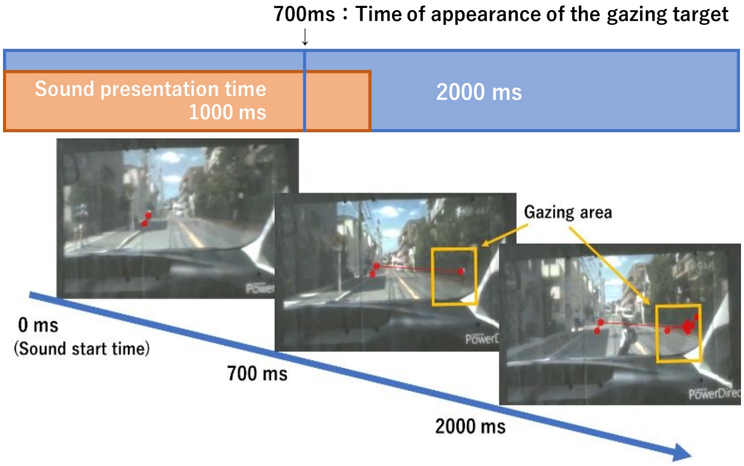


Figure 2: Analysis time line.

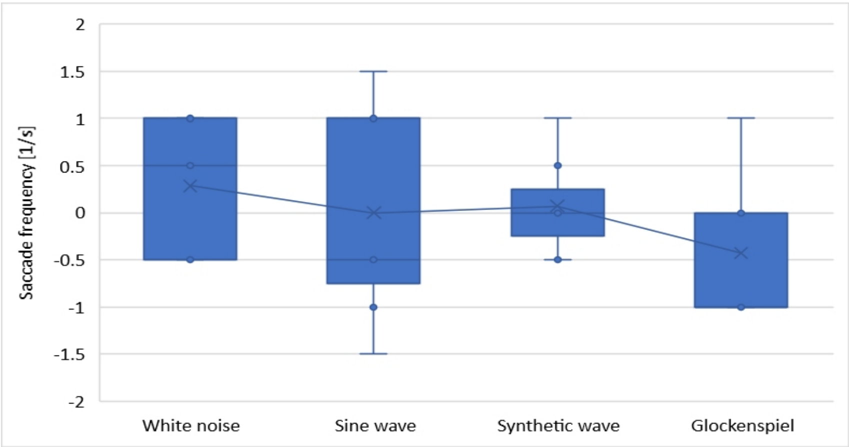
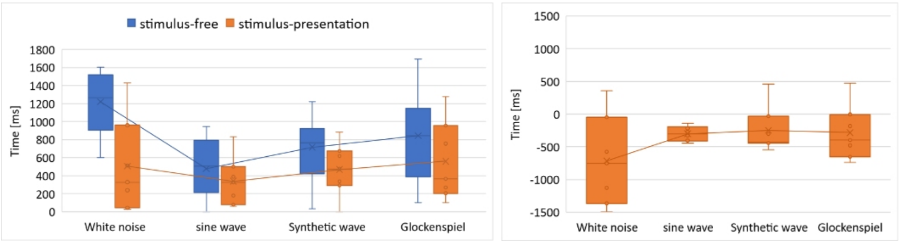


Figure 3: Difference in saccade frequency [⟨stimulus – presentation⟩ – ⟨stimulus – free⟩].



$*$ ··· $p < 0.05$

Figure 4: Induction time to the gaze region.

The plotted data of the fixation viewpoint was matched to the driving video because of finding the cause of these results. When white noise and sine waves were presented as stimulus sounds, it was found that the subjects' viewpoint was toward the red back lamps and the vehicle stopped in front of them. However, the other two conditions were not characterized by a concentration of viewpoints on some specific objects. This is considered to be the result of the comparatively small quantity of conspicuous objects in the presented driving video. From the above reasons, we argue that attention is focused on the conspicuous object or objects to which the attention is naturally directed during driving operations, that attention via visual stimuli takes precedence in gaze induction, and that auditory stimuli play a complementary role in guidance.

The two conditions of driving videos that did not show a reduction in guidance time to the gaze region have in common that there is less induction by visual information. Differences were observed in the guidance time to the gazing target, which could be caused by differences in induction effects among different sounds. As shown in Figure 5, there is a significant difference in the induction time to the gazing target in the condition where the synthetic wave is presented. The synthetic waves used in the experiments were considered to have good localization properties. We suggest that the synthesized waves may have enabled the subjects to focus more quickly on the gaze target, as the directional nature of the sound is an easily recognizable cue compared to instrumental sounds. In addition, John M. Findlay et al. stated in their book that spatial representation in the superior colliculus allows people to direct their gaze in the direction of the sound source (John, M. F. and Iain, D. G. (2006)).

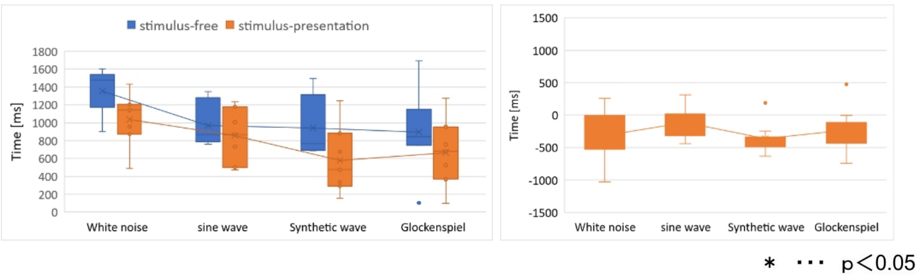


Figure 5: Guidance time to the gaze target.

Finally, the results of the post-experiment questionnaire are shown in Figure 6.

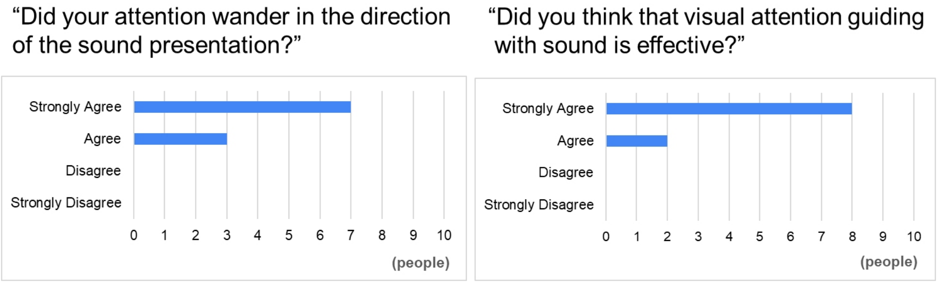


Figure 6: Questionnaire results.

However, more than half of the subjects responded in the questionnaire that they felt their attention was drawn in the direction of the sound. In addition, some subjects commented in the free description column that they could not find the target of their attention after the sound was heard. These results suggest that there may have been a problem with the timing of the preceding presentation of the stimulus sound.

CONCLUSION

It appears that the stimulus sound has little effect on the occurrence of saccades. We hypothesize that auditory stimuli might induce gaze. However, the effect tends to be stronger when an eye-catching object appears in front of the eyes, and it is not clear whether stimulus sound alone is sufficient to control the eye's gaze in detail. It is difficult to compare the differences in induction effects among the four different stimulus sounds since the results also suggest that visual information is preferred for attention in this study. To verify the effectiveness of the induction, it is necessary to measure the gaze of the subjects using the driving video with the same content. Based on the responses of the subjects to the questionnaire, we would like to investigate the future relationship between sound representation methods and features of visual stimuli and induction effects.

ACKNOWLEDGMENT

I would like to thank my deepest gratitude to Professor Yoshihisa Nakatoh for her guidance and encouragement in carrying out this research. I would also like to thank my sincere gratitude to Assistant Professor Yujie Li and Mr. Hiroki Uchida for their enthusiastic cooperation and guidance.

REFERENCES

Association for Electric Home Appliances. (2018). "*Kadenseihin niokeru Sousaseikoujou no tameno Houchion ni kansuru Gaidorain Dai2han [Kaitei]* (Guideline for sound alarms to improve operability in home appliances, 2nd edition [Revised])". Association for Electric Home Appliances.

Cabinet Office. (Jan. 23, 2023). "White Paper on Traffic Safety in Japan", the Cabinet Office home page, https://www8.cao.go.jp/koutu/taisaku/r03kou_haku/zenbun/genkyo/feature/feature_01.html

Douglas, P. Munoz, B. D. (1995) "evidence for interactions between target selection and visual fixation for saccade generation in humans", Exp Brain Res, 103, pp. 168–173.

Ito, K., Yamanaka, T., Goto, M., Miki, Y. (2014) "*Syasitunaionkyotokusei wo Kouryosita Houchion no Onnzoteiigijutu no Kentou* (Investigation of sound image localization technique for alarm sound considering cabin acoustic characteristics)", Transactions of the Society of Automotive Engineers of Japan, Inc., Vol. 45 No. 5, pp. 877–882.

JAF (Japan Automobile Federation). (Dec. 7, 2022) "'*Jisshaban' Kikennyochi.Jikokaihi Tore-ningu* (Live Risk Prediction Traning)". JAF Website: https://jaf.or.jp/common/safety-drive/online-training/risk-prediction

John, M. F. and Iain, D. G. (2006) Active vision: the psychology of looking and seeing, Honda, H. Kitaohji Shobo Co, Ltd.

Morikawa, D. Shimakura, N. Hirahara, T. (2009) "*Suiheimenonzoteii ni Hitsuyou na Syuuhasuutaiiki* (Frequency range required for horizontal plane sound image localization)", Proceedings of Acoustical Society of Japan, ASJ., Autumn term, pp. 2–3.