2009 International Workshop on Nonlinear Circuits and Signal Processing
NCSP'09, Waikiki, Hawaii, March 1-3, 2009

Effects from spatial cues on detectability of alarm signals in car environments

Naoki Kuroda[1], Junfeng Li[1], Yukio Iwaya[2], Masashi Unoki[1], and Masato Akagi[1]

1) School of Information Science,
Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa 923–1292 Japan
Phone/FAX:+81-761-51-1699(Ex.1391)/+81-761-51-1149
Email: {n-kuroda, junfeng, unoki, akagi}@jaist.ac.jp

2) Research Institute of Electrical Communication,
Tohoku University
2-1-1 Katahira, Aoba-ku, Sendai 980-8577 Japan
Phone/FAX:+81-22-217-5461 / +81-22-217-5535
Email: iwaya@riec.tohoku.ac.jp

Abstract

It is crucial to correctly detect alarm signals in noise conditions, for instance, especially in car environments. However, alarm signals are possibly masked by noises from engine, friction between tires and road, etc., in car environments. To design the alarm signals that can be easily detected, it is necessary to first understand the perceptual characteristics of alarm signals in noisy environments. In this paper, the masked thresholds of alarm signals in the presence of car noise were measured in the virtual acoustic environments generated by using head-related transfer functions (HRTFs). The main conclusions are: a) when the frequency component of alarm signals is 1.0 kHz, interaural time difference (ITD) and interaural phase difference (IPD) have great effect on the detectability of alarm signals in car noise; b) when the frequency component of alarm signals is 2.5 kHz, interaural level difference (ILD) except ITD and IPD also plays an important role in alarm signal detection when the signal is fixed in the front of listener.

1. Introduction

Alarm signals are sounds to provide informative notice to users. These signals convey a variety of information such as starting, ending, and alarm information to users. These signals are appropriately used for many purposes. It is important to detect them correctly in many sceneries. There are, however, cases where alarm signals cannot be correctly perceived in real environments due to noises. In car environments, for example, alarm signals are possibly masked by noises, and resulting in potential critical accidents in some cases. Therefore, the method for presenting alarm signals to be perceived accurately for drivers must be demanded. For this purpose, we must obtain fundamental data about perception of alarm signals in real environments, in particular, car environments.

For signal detection in the presence of noise, Ebata *et al.* reported that the detectability of signals can be improved by using directional information[1]. Moreover, in free sound field, Saberi *et al.* showed the detectability of signals was improved when the signal and masker were spatially separated[2]. This phenomenon is referred as to spatial release from masking (SRM). For the occurrence of SRM, ITD and ILD are regarded as the significant spatial cues. On the other hand, according to varying the frequency components which signals compose, binaural masked level difference (BMLD) occurs at the same time. For the occurrence of BMLD, IPD is regarded as the significant spatial cue. It has been confirmed to use both ITD and IPD as the spatial cues for perception of alarm signals in car noise[3][4]. The effectiveness of ILD has not yet considered in these conditions, although ILD has been known as one of the significant spatial cues for signal detection in free sound field.

In this study, the listening experiments were conducted to clarify the effectiveness of the spatial cues for perception of alarm signals in car environments. The experiments were carried out in the virtual anechoic environments by using HRTFs, which can simulated presentation of loud speakers in an anechoic room. As the advantages of using HRTFs, it is possible to extract ITD, IPD, and ILD respectively and easily control presentation directions of a signal.

2. Individualization of HRTF

HRTF is the acoustic transfer function from a sound source to an ear drum of a listener. Virtual sound presented to the listener from any point in three-dimentional space can be simulated by convoluting stimuli with impulse response of HRTFs. However, suitable HRTF is needed for each listener because HRTFs depend on physical characteristics of each listener. Otherwise, the ability in localizing sound images definitely decreases. Therefore, we must individualize HRTF for each listener to carry out listening experiments.

In this study, we find out the individualizing HRTF for each listener using a determination method of optimum impulse-response by sound orientation (DOMISO), which based on a tournament-style listening experiment[5]. DOMISO is the individualized method of HRTF which based on subjective evaluations by using tournament among HRTF data for listeners. 114 HRTF data which was obtained in the anechoic chamber in Tohoku university were used to individualize HRTFs for each listener.

Localization experiments of sound sources were carried out, in order to objectively evaluate whether the individualized HRTFs obtained by using DOMISO were suitable. The virtual sound sources were set 0° to 350° at the step size of

10°, and stimuli were randomly presented to listeners.

The evaluation criteria were set over 70% when the range of correct answers was ±10°, over 90% when the range of correct answers was ±20°, and under 10% in front-back confusion error. Only the listeners who satisfied the evaluation criteria participated in the listening experiments.

3. Listening experiments

It is important to clarify how spatial cues effect perception of alarm signals in noisy environments. As first step, we carried out the listening experiments under the same condition as that of Saberi *et al.*[2] (Experiment I). As second step, we carried out the listening experiments for perception of alarm signals in car noise environments (Experiment II).

3.1. Experiment I

3.1.1. Purpose

ITD and ILD have great effect on SRM. Saberi *et al.* carried out listening experiments by using loud speakers in an anechoic chamber. As their results, in free sound field, as the difference of presentation directions between the signal and noise increased, the amount of SRM was bigger and bigger. The maximal degree of masking release was about 14 dB[2]. For the occurrence of SRM, Nakanishi *et al.* showed that about 7 dB of masking release was caused by ITD in a sound-proof room[3]. Therefore, ITD is a significant spatial cue for perception of the pulse signals in white noise. However, the effectiveness of ILD has not yet considered. In order to clarify the effectiveness of ILD for the detectability of signals in noise, we carry out the experiments by using individualized HRTF. From the above-mentioned, the purposes of this experiment are :

1. To confirm whether the virtual acoustic environment generated by using individualized HRTF to the particular participant can simulate the presentation using loud speakers in an anechoic chamber.

2. To confirm how large degree ILD effects SRM in the simulated environment.

3.1.2. Methods

A pulse train signal, which was composed of 62.5-μs 100 rectangular pulses per 1 s, was used as the target signal. White noise was used as the masker. Stimuli consisted of a 1-s target signal and a 2-s masker. Here, the sampling frequency was 48 kHz.

Figure 1 shows the control method of arrival directions of the sounds by varying ITD which was calculated by

$$ITD = \frac{d}{c} = \frac{r(\theta + \sin\theta)}{c} \quad (1)$$

where r in [m] is the radius of head, θ in [rad] is the direction of the sound source, c in [m/s] is the sound velocity and d in [m] is the path difference from sound source to both ears. In this paper, r was set to 0.09 m and c was set to 343.5 m/s. Presentation directions of the target signal (or the masker) were varied from 0° to 90° at the step of 15°, the direction in front of the listeners was 0°. In this paper, the configurations of the target signal and the masker are assumed to be $S_m N_n$ (m=0, 15, •••, 90, n=0, 15, •••, 90). For example, $S_{60}N_0$ represents that the arrival directions of the target signal and masker are 60° and 0°, respectively.

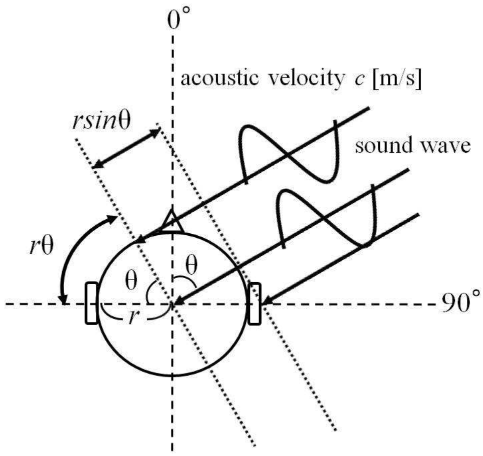

Figure 1: Illustration of difference in arrival time between both ears of a sound wave from a sound source in a direction of the angle θ radians for the observer.

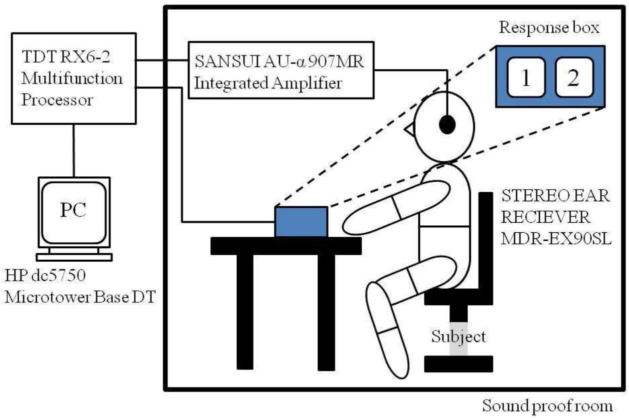

Figure 2: Environment for the masking experiments.

The arrival directions of the sound sources were varied by convoluting the stimuli with the corresponding HRTF for each direction when ITD and ILD were used as the spatial cues.

Figure 2 shows the diagram of the apparatus used in the experiment. The experiment was carried out in a sound-proof room by using Tucker-Davis Technologies (TDT) SystemIII. Stimuli were presented to each listener through ear phones (MDR-EX90SL).

3.1.3. Procedure

In this experiment, the masked thresholds of the target signals by the masker were measured by using the method of limits. There are descending and ascending series in the method of limits. In descending series, at the beginning of the experiment, the sound pressure level of the target signal in the stimuli was chosen randomly from the range that the listener can distinctly detect the target signal. Then, the sound pressure level of the target signal was varied from high to low at the step of 1 dB. In ascending series, at the beginning of the experiment, the sound pressure level of the target signal in the stimuli was chosen randomly from the range that the listener cannot distinctly detect the target signal. Then, the sound pressure level of the target signal was varied from low to high at the step of 1 dB. In addition, the starting position of the target signal in the stimuli was chosen randomly. The sound pressure level of the masker was fixed to 65 dB at the beginning of the experiments. Both descending series and ascending series were carried out 10-trials respectively. When the difference in the mean of each series was 2 dB or less, the masked threshold was determined as the mean of all measurements.

Six graduate students aged from 23 to 26, five males and one female, participated in this experiment. All had normal hearing (15 dB HL from 0.125 to 8 kHz) and experiences in attending the other psychoacoustical experiments.

3.1.4. Results and Discussion

Figure 3 shows the results of the mean masked thresholds at each direction for perception of the pulse train signals in white noise. The vertical axis shows the relative masked thresholds which are normalized by the masked thresholds at S_0N_0. The horizontal axis shows azimuth of the pulse train signal or the white noise. The blue lines show the results under the conditions that ITD was used as the spatial cue, and the red lines show the results under the environment by using individualized HRTF. In addition, the solid lines show the results of the S_mN_0 (m=0, 15, •••, 90), and the dotted-line show the results of the S_0N_m (m=0, 15, •••, 90). The error bars show standard deviation of relative masked thresholds for each presentation direction. These results showed that almost the same tendency observed as the previous research[3], when the listeners used ITD as the significant spatial cue. In the S_mN_0, ITD is used as the significant spatial cue because the blue and red lines are the same tendency. There is the range that the masked thresholds in the red line are higher than that in the blue line. This means that the effectiveness of ITD decreases by adding the effectiveness of ILD, resulting in proceeding the interaural differences becomes difficult, and the masked thresholds slightly increase. In the S_0N_m, the maximum degree of masking release is about 16 dB in the virtual acoustic environment generated by using individualized HRTF, and this almost equals to that of Saberi *et al.*[2]. From the above-mentioned results, we found that the experiments by using individualized HRTF can simulate the experiments by using loud speakers in an anechoic chamber and ILD has great effects.

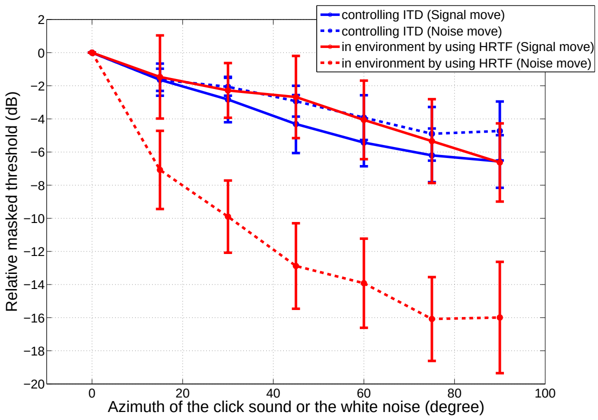

Figure 3: The mean masked thresholds as function of the presentation direction of the signal or masker for perception of the pulse signal in white noise.

3.2. Experiment II

3.2.1. Purpose

In the previous papers, the effectiveness of ITD and IPD for perception of alarm signals in car noise, was investigated[4]. As the results, when IPD of the signal or noise was homophasic, the detected limit of the alarm signals increased, in contrast, when IPD of the signal or noise was antiphasic, the detected limit of the alarm signals decreased. From these results, it was important to use both ITD and IPD as the spatial cues for perception of the alarm signals in car environments. However, the effectiveness of ILD has not yet been considered. Therefore, the purpose of this experiment is to investigate how ILD effects the detectability of the alarm signals in car noise.

3.2.2. Method

In this experiment, the target signal and masker were the alarm signals and car noise respectively. Alarm signals were the signal conveyed most warning provided JIS S 0013[6]. These signals had the repeated patterns of ON and OFF (ON = 0.1 s, OFF = 0.05 s) for 1 s. The frequency components of the alarm signals were 1.0 and 2.5 kHz. The car noises were recorded in a car interior with the window open while the car was traveling at 60 km/h. The sampling frequency was 48 kHz.

The setting method of ITD and the environment of this experiment are the same as those in Experiment I.

3.2.3. Procedure

Procedure was the same as in Experiment I.

3.2.4. *Results and Discussion*

Figure 4 shows the results of the mean masked thresholds at each direction for perception of the alarm signals in car noise. When the alarm signal or car noise was provided the listeners to ITD as the significant spatial cue, we observed that ITD and IPD effect the detectability of the alarm signals in car noise. These results showed that almost the same tendency in the previous results[4] was observed. However, there is difference between the amount of masking release at the S_mN_0 and S_0N_m. This means that the extent of effectiveness of ITD and IPD is different between these conditions. For the virtual acoustic environments generated by using individualized HRTF, the amount of masking release was smaller than that of the previous studies in the $S_mN_0 condition$[4]. We reconsider that the calculation of interaural correlation becomes difficult due to the effect of ILD. On the other hand, interaural correlation could be calculated in the S_0N_m because the sound pressure level of the alarm signals equals in both ears. In the S_0N_m, the amount of masking release for 1.0 kHz is almost the same tendency as in the virtual acoustic environment generated by using individualized HRTF. This means that the listeners did not use ILD as the significant cue. In contrast, amount of masking release for 2.5 kHz signal is greater than that of the previous study[4]. Hence, the detectability of the alarm signals in car noise is improved by using ILD as the significant cue, and we can then observe the occurrence of BMLD by using IPD. In a summary, when frequency component of the alarm signal is 2.5 kHz, we clarified that ILD has great effects for the perception of the alarm signals in car noise at the S_0N_m.

4. Conclusions

In this study, we measured the masked thresholds of the alarm signals in the virtual acoustic environment generated by using individualized HRTF. As a result, when the frequency component of the alarm signal was 1.0 kHz, we did not observe that ILD effects SRM in both the S_mN_0 and S_0N_m. While, we found that the effectiveness of ITD and IPD was observed in these conditions. When the frequency component of the alarm signal was 2.5 kHz, we observed that ILD had great effects SRM in the S_0N_m. As the summary, in this paper, we clarified the followings.

1. When the frequency component of the alarm signals was 1.0 kHz, ITD and IPD were used as the significant spatial cues for perception of the alarm signals in car noise.
2. When the frequency component of the alarm signals was 2.5 kHz, ITD and IPD were used in the S_mN_0, ITD, IPD, and ILD are used as the significant spatial cues for perception of the alarm signals in car noise.

References

[1] M. Ebata, T. Sone, and T. Nimura, "Improvement of hearing ability by directional information," *J. Acoust. Soc. Am.*, vol. 43, no. 2, pp. 289–297, 1968.

[2] K., Saberi, Dostal, L., Sadralodabei, T., Bull, V., and Perrot, D., R., "Free-field release from masking," *J. Acoust. Soc. Am.*, vol. 90, no. 3, pp. 1355–1370, 1991.

[3] J., Nakanishi, M., Unoki, M., Akagi, "Effect of ITD and Component Frequencies on Perception of Alarm Signals in Noisy Environments," *J. Signal Processing*, vol. 10, no. 4, pp. 231–234, 2006.

[4] H., Uchiyama, M., Unoki, and M., Akagi, "Improvement in detectability of alarm signal in noisy environments by utilizing spatial cues," *Proc. WASPAA2007*, pp. 74–77, New Paltz, NY, 2007.

[5] Y. Iwaya, "Individualization of head-related transfer functions with tournament-style listening test: Listening with other's ears," *Acoust. Sci. & Tech.*, vol. 27, no 6, pp. 340–343, 2006.

[6] JIS S 0013, Guidelines for the elderly and people with disabilities – Auditory signals on consumer products., 2002.

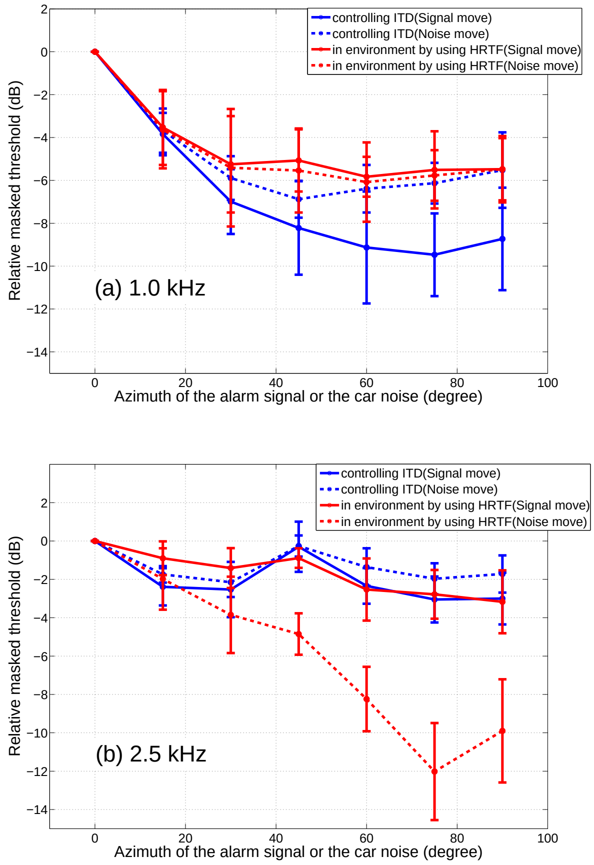

Figure 4: The mean masked thresholds as function of presentation directions of the signal or masker for perception of the alarm signal in car noise : (a) 1.0 kHz and (b) 2.5 kHz.