Title	Improvement of Detectability of Alarm Signal in Noisy Environments by Utilizing Spatial Cues
Author(s)	Uchiyama, Hideaki; Unoki, Masashi; Akagi, Masato
Citation	2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics: 74-77
Issue Date	2007-10
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/7802
Rights	Copyright (C) 2007 IEEE. Reprinted from 2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of JAIST's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.
Description	

IMPROVEMENT IN DETECTABILITY OF ALARM SIGNALS IN NOISY ENVIRONMENTS BY UTILIZING SPATIAL CUES

Hideaki Uchiyama, Masashi Unoki, and Masato Akagi

School of Information Science, Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa, 923-1292 Japan
`{h-uchiya,unoki,akagi}@jaist.ac.jp`

ABSTRACT

We measured how well alarm signals could be detected in the presence of car noise as a function of interaural time difference (ITD) and interaural phase difference (IPD). Alarm signals intended to warm people of dangerous situations need to be perceived accurately in real environments because noise in a real environment can mask the alarm signals and make them undetectable. Pulse train signals and five alarm signals were used to check whether the spatial release from masking (SRM) occurred. The results showed that SRM occurred for all signals and that alarm signal detectability could be improved by utilizing not only ITD but also IPD of the signal. This effect depended on the relationship between ITD and IPD. In addition, ITD and IPD of the arrival direction difference of the alarm signal in the masker greatly influenced occurrence of SRM: this could be interpreted as binaural masking level difference (BMLD). These results suggest that spatial cues of the arrival direction of an alarm signal in comparison with the masker direction have to be considered in conveying warnings accurately and efficiently without loss of information.

1. INTRODUCTION

Audible alarm signals are used to attract the attention of persons in many everyday activities, for example, the beeps and/or melodic sounds of electronic products such as washing machines and effective sounds in car navigation systems [1]. Therefore, these signals need to be perceived accurately and efficiently by everyone. For this purpose, alarm signals with many different stimulus shapes have been studied to check if they are perceived adequately, e.g., by Mizunami *et al.* [2]. Although alarm signals should be perceived correctly even in real environments so that the intended person knows when and what events have occurred, there has been less study on the robustness of alarm signal detectability. Interference is produced through the masking effects of noise in noisy environments, and this can dramatically reduce alarm signal detectability. This is a serious problem because it leads to many dangerous situations for persons who fail to hear important alarm signals. Therefore, it is important to present alarm signals in such a way that they can be correctly detected in any environment.

On the other hand, Ebata *et al.* reported that the ability to detect a signal sound in the presence of noise could be improved by utilizing directional information [3]. Saberi *et al.* also reported that the detectability of a pulse train signal against white noise in the free field could be improved when the signal and masker were spatially separated [4]. This means that the masking threshold (detectability) can be improved using spatial cues in binaural hearing. Therefore, this is referred to as “spatial release from masking (SRM)” [4]. If SRM occurs for alarm signals in a noisy environment, it can suppress the influence of the masking effect produced by noise on alarm signals, which facilitate the perception of their existence and directions, so it can help us in designing the way in which alarm signals are presented.

It is well known that interaural time difference (ITD) and interaural level difference (ILD) are used as significant cues in SRM [5]. The aims of our work are to confirm that SRM occurs for alarm signals in noisy environments and then to determine whether SRM can be accounted for by these spatial cues. However, in the above-mentioned experiments, SRM were measured in the case where signal and noise were presented through loudspeakers in the free field as a function of the direction of either the signal or noise with respect to the subject. It is difficult to investigate SRM for alarm signals in the presence of noise without removing the influence of reverberation and background noise. It is also difficult to investigate the separate influences of ITD and ILD in SRM.

As the first step toward investigating the detectability of binaural alarm signals in SRM, we first scaled down the experiments in a free field (loudspeaker presentation) to experiments in a soundproof room (headphone presentation) in terms of experimental design to cancel out the effects of the surrounding background noise and reverberation and to control spatial cues (ITD and ILD) separately. ITD was used as the dominant cue in our experiments to investigate the detectability of alarm signals in noisy environments because it can be more easily controlled via the azimuth of stimuli than ILD. The masking thresholds of binaural alarm signals in the presence of noise were measured as a function of ITD induced by an alarm signal by varying the component-frequency [7]. As a result, it was found that SRM occurred not only for the pulse train signals but also for alarm signals as ITD increased. It was also found that another spatial cue, interaural phase difference (IPD), improved the detectability of alarm signals against white noise depending on the relationship between signal frequency and IPD.

As the second step, we checked whether SRM occurred for an alarm signal in the presence of realistic noise (using car noise instead of white noise). We also investigated ways of improving detectability of alarm signals in the environments by considering spatial cues (ITD and IPD) and we determined how much SRM was influenced by the relationship between signal frequency and spatial cues (ITD and IPD).

2. PREVIOUS WORK AND DESIGN OF EXPERIMENTS

In our previous work [7], we conducted two experiments using pulse train signals and three alarm signals (1.5, 2.0, and 2.5 kHz) to determine the detectability of alarm signals in the presence of

978-1-4244-1619-6/07/$25.00 ©2007 IEEE

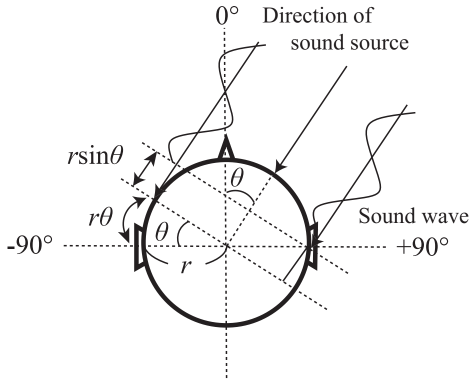

Figure 1: Illustration of the difference in arrival times two ears of a sound wave from a sound source in a direction angle θ radians from the observer.

white noise. In the first experiment, the same as in the experiment of Saberi *et al.*, masking thresholds of pulse train signals in the presence of white noise were measured as a function of ITD. The results showed that SRM occurred in these conditions and its maximum level of masking release (decresed level of masking threshold) was about 8 dB. In the second experiment, masking thresholds of alarm signals (1.5, 2.0, and 2.5 kHz) in the noise were measured as a function of ITD. Results showed that SRM also occurred for these signals, but the masking thresholds were influenced not only by their ITD but also their IPD. This depends on the component-frequency of the alarm signal. Maximum levels of SRM were about 2 dB at 1.5 and 2.0 kHz and about 3 dB at 2.5 kHz. From these results, we conclude that there are two issues: (1) whether SRM occurs for alarm signals in the presence of real noise or not and (2) how can SRM be accounted for by the relationship between the spatial cues (ITD and IPD).

In this paper, we report on two experiments carried out to investigate these two issues. The aim of first experiment was to clarify whether the difference between two noises (white noise and realistic noises) influences the occurrence of SRM. Thus, the detectability of pulse train signals in realistic noise was measured as a function of ITD. The aim of the second experiment was to investigate how the detectability of alarm signals amid realistic noise could be improved when the given ITD and IPD were varied. Thus, the detectability of five alarm signals (1.0, 1.5, 2.0, 2.5, and 4.0 kHz) was measured as a function of the relationship between signal frequency and spatial cues (ITD and IPD). Their detectability was also measured as a function of the ITD of the alarm signal for a fixed arrival direction of noise and vice versa.

3. EXPERIMENT I

3.1. Purpose

This experiment had three purposes: (1) to confirm that SRM occurs for alarm signals in a realistic noise condition by using only the ITD cue, (2) to examine whether perceptual characteristics are similar to our previous results [7], and (3) to investigate whether a different tendency is observed for the ITD of the signal for a fixed noise direction and vice versa.

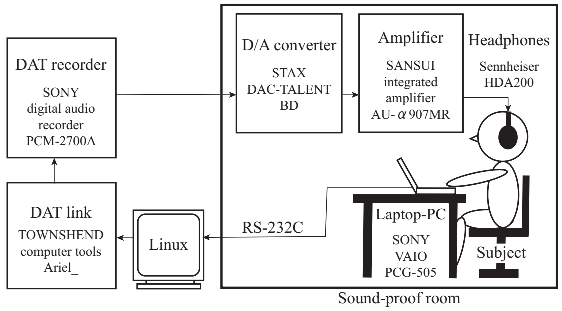

Figure 2: Environment of the signal detection experiments.

3.2. Method

The pulse signal was composed of 62.5 μs rectangular pulses presented at a rate of 100 pulses per second. This is the same as the signal that Saberi *et al.* used [4]. A 1-s pulse train signal was used as the target signal in this experiment. As a masker, we used 2 seconds of car noise, which was recorded in a car interior with the window open while the car was traveling at 60 km/h. Here, the sampling frequency was 48 kHz.

Figure 1 shows how we controlled the direction of the sound source (θ in radians) as a function of ITD. Here, ITD can be described as

$$\begin{aligned} \mathrm{ITD} &= \frac{d}{c} = \frac{r}{c}(\theta + \sin\theta), && (1) \\ d &= r\theta + r\sin\theta, && (2) \end{aligned}$$

where r in meters is the radius of a human head, θ in radians ($-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$) is the direction of the sound source, c in m/s is the sound velocity, and d in meters is the route difference between the two ears. Here, r was 0.09 m and the sound velocity was 343.5 m/s. θ was set to $\frac{\pi}{12}$, $\frac{\pi}{6}$, $\frac{\pi}{4}$, $\frac{\pi}{3}$, $\frac{5\pi}{12}$, and $\frac{\pi}{2}$ in radians ($15^\circ, 30^\circ, 45^\circ, 60^\circ, 75^\circ$, and 90°) representing the signal (or the noise) source moving to the front-right of the subject when the median plane was assumed to be 0°. The signal-noise configurations were as follows. The stimulation presentation condition in which the ITD of the signal was varied while the noise direction was fixed (0°) was assumed to be $\mathrm{S}_m\mathrm{N}_0$ ($m = 0, 15, \cdots, 90$). The opposite presentation condition was assumed to be $\mathrm{S}_0\mathrm{N}_m$ ($m = 0, 15, \cdots, 90$). For example, $\mathrm{S}_0\mathrm{N}_0$ denotes the condition in which both the signal and the noise were located at 0°.

Figure 2 shows a schematic diagram of the apparatus used in the experiment. The experiment was carried out in a sound proof room. Stimuli were presented to each subject through headphones (Sennheiser HDA200) while the subject was operating an application running on a laptop PC (Sony VAIO PCG-505). This application establishes serial communication with a PC-Linux system to present stimuli in order to collect the subject's responses.

3.3. Procedure

The method of limits was used to measure masking thresholds (detectability of the pulse train signals) in this experiment. This method includes descending and ascending series. In the descending series, when the experiment began, the sound pressure level of the signal in the stimuli was chosen randomly from the range for

978-1-4244-1619-6/07/$25.00 ©2007 IEEE

which the subject could distinctly perceive the signal. The signal level was varied from high to low with a step size of 1 dB. In the ascending series, at the beginning of experiment, the initial signal level was set in the range where the subject could not distinctly perceive the signal. The signal level was then varied from low to high with a step size of 1 dB. The stimulus was presented via the system shown in Fig. 2 as a 1-s signal in a 2-s noise. The signal position was set at random close to the central position in noise where the subjects could not detect the signal using the onset cue. Here, the masker was presented at 65 dB SPL. The subject was required to detect the signal in the presence of noise. Ten trials were carried out for each series, descending and ascending. When the difference in the mean of each series was 2 dB or less, the masking threshold was determined as the mean of all measurements. A run was repeated until the difference between any two of the averaged thresholds in two ascending series was within 2 dB.

Seven graduate students aged between 23 and 26, six male and one female, participated in this experiment. All had normal hearing (15 dB HL or less for both ears from 0.125 to 8 kHz) and had experience participating in other psychoacoustical experiments.

3.4. Results and discussion

The mean masking thresholds obtained for the pulse train signal are shown in Fig. 3(a). All thresholds as detectability are plotted relative to $\mathrm{S_0N_0}$. Error bars represent the standard deviation. The blue dotted line indicates the detectability for $\mathrm{S}_m\mathrm{N}_0$, and the red broken line indicates the detectability for $\mathrm{S_0N}_m$. As the signal source was moved to the front-right of the subject (0° to 90°), masking thresholds decreased as a function of azimuth.

These results exhibit almost the same tendency as the previous results [7]. The only difference was the maximum level of SRM for different types of noise (white noise or car noise). While masking thresholds were decreased by about 8 dB in the condition of $\mathrm{S_{90}N_0}$ with a white noise, the masking thresholds measured in this experiment were decreased by about 2 dB for $\mathrm{S_{90}N_0}$ with car noise. We considered the reason for this difference. First, the power spectra of the two types of noise that we used were different: constant (white) power spectrum for white noise and mainly descending slope (pink) power spectrum for car noise. Second, the power spectrum of the pulse train signal was diffused, so it was spectrally spread. Therefore, it is possible that subjects were detecting the spectral spreads at higher frequencies. Hence, it may be considered that the amount of masking release tended to be lower than in the results of the previous experiment [7].

The results for both conditions ($\mathrm{S}_m\mathrm{N}_0$ and $\mathrm{S_0N}_m$) were almost the same. This is consistent with the previous results [4] .

4. EXPERIMENT II

4.1. Purpose

In this experiment, alarm signals with widely different component-frequencies in the presence of car noise were used as stimuli. The alarm signals were set in different arrival directions as a function of ITD. However, in this case, it was predicted that interaural phase difference (IPD) could be used as a cue for detecting the signal in the noise, depending upon the signal frequency in the ITD condition. Thus, to investigate how the SRM characteristics could be accounted for when ITD and IPD were given as cues, masking thresholds of alarm signals in realistic noise were measured as a function of ITD induced by the alarm signals.

Table 1: Relationship between ITD and IPD.

θ (°)	ITD (ms)	IPD ($\pi/2$ rad)	IPD (π rad)
15	0.136	2.5, 4.0 kHz	—
30	0.268	2.0 kHz	4.0 kHz
45	0.391	1.5 kHz, 4.0 kHz	2.5 kHz
60	0.501	1.0 kHz	2.0 kHz, 4.0 kHz
75	0.596	2.5, 4.0 kHz	—
90	0.674	2.0 kHz	1.5 kHz, 4.0 kHz

4.2. Method and procedure

It is well known that the ability of localize a sinusoidal signal by using the ITD is reduced when the signal frequency is over about 1.5 kHz and the ability of IPD-based localization is also reduced when the frequency is less than 1.8 kHz [6]. Above these frequencies, it is known that ITD is used in the temporal envelope and IPD is useful for localization over 1.8 kHz. In the previous experiment, three alarm signals (1.5, 2.0, and 2.5 kHz) were used [7]. In order to investigate SRM via spatial cues (ITD and IPD), five alarm signals (1.0, 1.5, 2.0, 2.5, and 4.0 kHz) were used as signals. The alarm signals were composed of signals that conveyed most warnings provided JIS S 0013 [1]. These signals were repeated patterns of ON and OFF (ON = 0.1 s, OFF = 0.05 s) for 1 s. The frequencies of the components of the alarm signals were 1.0, 1.5, 2.0, 2.5, and 4.0 kHz. ITDs were set to the same values as in Experiment I. Car noise for 2 s was also used as a masker and its level was also set to be 65 dB SPL.

4.3. Results and discussion

The mean masking thresholds obtained for the five alarm signals in the presence of car noise are shown in Figs. 3(b)-3(f). We found that SRM occurred for all alarm signals as a function of ITD. We also found that SRM for 1.5, 2.0, and 2.5 kHz alarm signals exhibited a similar tendency to the previous results [7]. However, the improvement in alarm signal detectability did not tend to decrease monotonically as ITD increased (moved from 0° to 90°). The SRM seemed to be V- or W-shaped. These features were seen in the previous results (1.5, 2.0, and 2.5 kHz) [7]. The maximum level of masking release at low frequency (8 dB at 1.0 kHz) was greater than that at higher frequency (about 2 dB at 4.0 kHz).

It may be possible to account for these results using the binaural masking level difference (BMLD) [6]. BMLD is well known as the level difference of the masking threshold ML_m dB when homophasic white noise and a pure tone are mixed and presented and as the masking threshold ML_π dB when homophasic white noise and an antiphasic pure tone are mixed and presented in both ears.

The results for the 1.5-kHz alarm signal are considered as a representative example. The duration of one period of 1.5-kHz sinusoidal wave is about 0.67 ms and the duration of a half period of this sinusoidal wave is about 0.335 ms. A time difference of 0.67 ms is the same as the ITD value for which SRM did not occur in the $\mathrm{S_{90}N_0}$ and $\mathrm{S_0N_{90}}$ conditions. In other words, in the $\mathrm{S_{90}N_0}$ and $\mathrm{S_0N_{90}}$ conditions, the alarm signal was presented to subjects with a one-period delay between the two ears. Therefore, it can be interpreted that a higher masking threshold was obtained at the $\mathrm{S_{90}N_0}$ and $\mathrm{S_0N_{90}}$ conditions related to IPD (π radians) because it became extremely close to the condition in which the homophasic alarm signal was presented to both ears. Similarly, a time differ-

978-1-4244-1619-6/07/$25.00 ©2007 IEEE

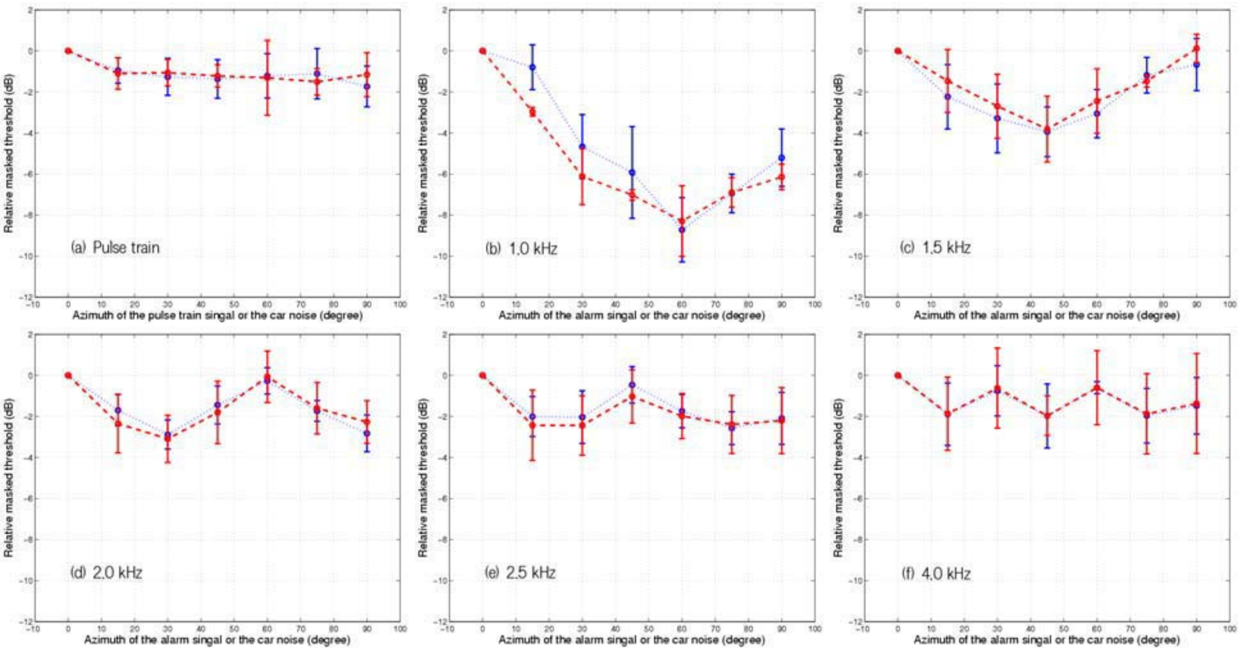

Figure 3: Mean masking thresholds for: (a) pulse train signal and for alarm signals at (b) 1.0 kHz, (c) 1.5 kHz, (d) 2.0 kHz, (e) 2.5 kHz, and (f) 4.0 kHz. All thresholds are plotted relative to the threshold for the $\mathrm{S}_0\mathrm{N}_0$. Filled circles and error bars show mean value and standard deviation of thresholds, respectively. The blue dotted line indicates results for $\mathrm{S}_m\mathrm{N}_0$ and the red broken line indicates results for $\mathrm{S}_0\mathrm{N}_m$.

ence of 0.335 ms is the same ITD value for which SRM occurred in the $\mathrm{S}_{45}\mathrm{N}_0$ and $\mathrm{S}_0\mathrm{N}_{45}$ conditions. It can be interpreted that the lower masking threshold was obtained in these conditions because the phase of the alarm signal was shifted by half a period ($\pi/2$) between the two ears.

For the other alarm signals (1.0, 2.0, 2.5, and 4.0 kHz), SRM can be accounted for by a function of the relationship between ITD and IPD corresponding to BMLD. This relationship is shown in Table 1. The amount of masking release can similarly be accounted for by interpreting it as BMLD. The value of BMLD is 15 dB at most at low frequencies (about 0.5 kHz) and about 2 – 3 dB at frequencies above 1.5 kHz [6]. Moreover, it occurs when the signal and the masker are spatially separated, which is the condition in which BMLD occurs in a real environment.

5. CONCLUSION

Two experiments were carried out using a pulse train signals and five alarm signals to investigate whether SRM occurred and how detectability can be improved by utilizing spatial cues (ITD and IPD). Experiment I showed that SRM occurred as a function of ITD in a realistic noise environment, but the amount of masking release was small. Experiment II showed that SRM for an alarm signal in the presence of realistic noise was influenced not only by the ITD but also by IPD of of the signal, depending on the signal frequency. In particular, we found that the lack of improvement in alarm signal detectability for conditions $\mathrm{S}_m\mathrm{N}_0$ and $\mathrm{S}_0\mathrm{N}_m$ where m corresponds to the IPD of π-phase and the improvement for these conditions where m corresponds to the IPD of $\pi/2$-phase were similar tendencies because BMLD with IPD is also related to SRM. Hence, we conclude that a spatial cue, IPD, also strongly affects the improvement in detectability, i.e., BMLD affects SRM.

6. REFERENCES

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

[2] T. Mizunami, K. Kurakata, H. Shimosako, and K. Matsushita, “Further examination of ON/OFF temporal patterns of auditory signals (completion signals and attention signal) recommended in JIS S 0013,” *Japanese Journal of Ergonomics*, **40**(5), 264-271, 2005.

[3] M. Ebata, T. Sone, and T. Nimura, “Improvement of hearing ability by directional information,” *J. Acoust. Soc. Am.*, **43**, 289–297, 1968.

[4] K. Saberi, L. Dostal, T. Sadralodabai, V. Bull, and D. R. Perrot, “Free-field release from masking,” *J. Acoust. Soc. Am.*, **90**(3), 1355-1370, 1991.

[5] C. Lane, N. Kopco, B. Delgutte, B. G. Shinn-Cunningham, and H. S. Colburn, “A cat’s cocktail party: Psychophysical, neurophysiological, and computational studies of spatial release from masking,” *Proc. ISH 2003*, 341-347, 2003.

[6] B. C. J. Moore, An Introduction to the Psychology of Hearing, Fifth Edition, Chap. 7, Academic Press, London, 2003.

[7] J. Nakanishi, M. Unoki, and M. Akagi, “Effect of ITD and Component Frequencies on Perception of Alarm Signals in Noisy Environments,” Journal of Signal Processing, **10**(4), pp. 231-234, July 2006.