Study on detectability of target signal by utilizing differences between movements in temporal envelopes of target and background signals

Yuta Yano, Ryota Miyauchi, Masashi Unoki, and Masato Akagi

School of Information Science, Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa, 923-1292 Japan
Phone/FAX:+81-761-51-1391/+81-761-51-1149
Email: {s0810064, ryota, unoki, akagi}@jaist.ac.jp

Abstract

In this paper, we investigated whether difference between movements in temporal envelopes of the target and background signals can improve the detectability of target signal against background signals. The movement of the signal was defined as the mean value of slopes as the first-order approximation of temporal envelope. Stimuli were generated by multiplying one of three processed temporal envelopes (movements of 88, 700, and 1313 mV/s) with one of three different noise carriers (1/2-octave bandwidth at the center frequencies (CFs) of 200, 525, and 1380 Hz). The auditory search experiment was carried out. Results showed that the detectability of target signal, that is, d', is improved as the difference of the movements increases, in which the CF of target signal was the same as that of the background signal. This tendency constantly appeared when the movement of the target signal was faster than that of the background signal. These results suggest that greater difference of the movements improves the detectability of the target signal.

1. Introduction

Audible alarm signals are used to attract the attention of persons in many everyday activities, for example, the beeps and melodic sounds of electronic products for providing starting and ending, and fire alarm sounds for emergency [1]. Therefore, these signals must 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]. There are, however, cases where alarm signals cannot be correctly perceived in real environments because they are masked or partially masked by background noise, so that the intended person knows when and what events have occurred. Therefore, it is important to present alarm signals in such a way that they can be correctly detected in any environment.

It is well known that the detectability of a signal 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 detectability can be improved by 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 noisy environments, 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.

On the other hand, a human can easily selectively listen to a target sound in real noisy environments that simultaneously contains various kind of sound such as conversation speech and environmental sounds. It has been considered that we are not attend to listen all signals in real environments but we are attend to listen a specific signal such as alarm signal in as the all signals. In the related study, there is what is called the "cocktail party effect" that is ability of human auditory system [5]. This ability is what we can perceive one sound in complex sounds. Consequently, it may be considered that human can perceive a specific alarm signal easily in noisy environments, by utilizing this ability.

In the most related study, there is, for example, auditory search tasks, studied by Asemi *et al.* [6]. They studied whether the detectability of the target against background signal can be improved, by carrying out auditory search experiments. In their experiments, narrow-band noise and pure tone were used and temporal fluctuations of the target was controlled as a cue to investigate improvement of the detectability of the target signal. Their results conducted that the temporal fluctuation of the target against background signals plays a role of the improvement in the detectability.

Based on their results, Kusaba *et al.* have investigated how much the temporal fluctuation between the target and background signal can improve the detectability of the target signal in auditory search experiments [7]. Their results suggest that lower similarity of temporal fluctuations between target and background signals can improve the detectability of the target in noisy environments. However, In their experiments, they used similarity of temporal fluctuations between target and background signals to account for the improvement of detectability. Thus, it is still unclear what characteristics of the movements that can contribute to the improvement are.

In this study, we have the same approach that Kusaba *et al.* had, to investigate the detectability of the target against background signal as the auditory search problem. Instead of the use of similarity between target and background signals, we reconsidered whether difference between movements in temporal envelopes of the target and background signals can improve the detectability of the target signal. This study aims to reveal what movements in temporal envelopes of the target signal can improve the detectability of the target signal.

2. Definition and control of the movement of signal

We defined the movement in temporal envelope of a signal as the mean value of slopes as the first-order approximation of temporal envelope.

$$v = \frac{A_0}{N-1}\sum_{n=2}^{N}\frac{|P_n - P_{n-1}|}{T_n - T_{n-1}} \tag{1}$$

where P_n is the values of peak/dip in temporal envelopes, T_n is the time, and N is the number of peaks/dips. A_0 is output reference-voltage (through the TDT SystemIII and the SANSUI AU-α 907MR Integrated Amplifier). The output reference-voltage to computerize amplitude value of 1.0 is 88 mV. Therefore, v is in mV/s.

Figure 1 shows an example of the temporal envelopes to derive their movements. Red line indicates a temporal envelope of a signal. Green line indicates the mean value of slopes of the temporal envelope. In this case, the number of peaks/dips is seven. Therefore, this movement in temporal envelope is 176 ($= 88/(7-1) \times 12$) mV/s.

We systematically controlled the movements, defined in the above by bandpass (BP) filtering on the modulation spectrum as follows. We assumed that signal can be expressed by multiplying temporal envelope with carrier. Here, band-noise was used as the carrier.

Figure 2 shows an example of stimulus generation, used in auditory search experiment. White noise is used to generate stimulus, as shown in Fig. 2(a). Figure 2(b) shows the modulation spectrum of (a). Figure 2(c) shows characteristics of the BP filter in the modulation frequency. This filter has the bandwidth of 2.5 Hz at the CF of 2 Hz. By BP filtering on the modulation spectrum, the processed modulation spectrum of them was obtained as shown in Fig. 2(d). The processed amplitude envelope was then obtained from the inverse Fourier transform of them as shown in Fig. 2(e). This was normalized by the averaged value of amplitude to be ± 0.5. Stimuli were generated by multiplying the temporal envelopes with the carriers. The rising and falling edge of the stimuli were created as 30-ms raised-cosine.

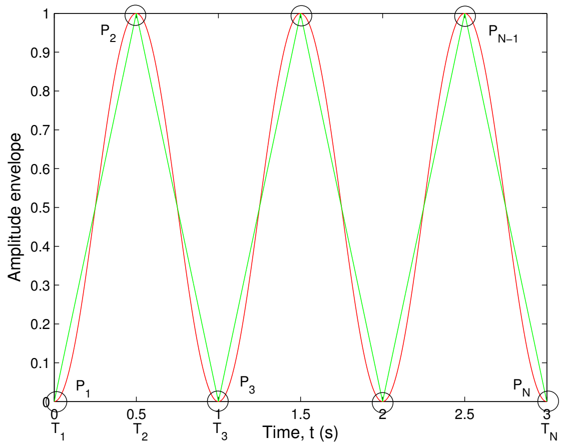

Figure 1: Example of the movement in the temporal envelope (176 mV/s).

To control the movements in temporal envelopes, we systematically control envelopes shown in Fig. 2(f) by using the Eq. (1). The movements tend to increase as the CF of filter increases and/or bandwidth of them widens. As the above, the three types of movements were generated by systematically controlling the modulation spectrum of them using BP filter.

3. Experiment

3.1. Apparatus

Figure 3 shows the diagram of the apparatus used in the experiment. The experiment was carried out in a sound-proof room. Stimuli were generated by Tucker-Davis Technologies (TDT) System III and presented to each subject through an amplifier (SANSUI, AU-α) and headphones (Sennheiser HDA200). A subject's response was recorded through the response box.

3.2. Stimuli

The movements in temporal envelopes we used in the experiment were 88, 700, and 1313 mV/s. The carrier was $1/2$-octave band-noise. The CF of the carrier were 200, 525, and 1380 Hz. Stimuli were composed of the three temporal envelopes with the three noise carriers. All stimuli have nine conditions.

3.3. Procedure

In auditory search experiment, there were two kinds of judgments: positive and negative judgments. In the positive judgments, two of the nine stimuli as the target and background signals were simultaneously presented to the partici-

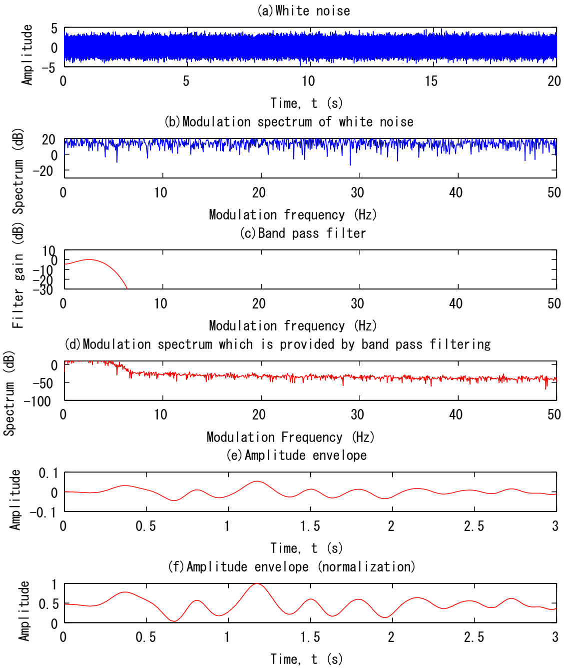

Figure 2: Examples of stimulus generation: (a) white noise, (b) modulation spectrum of (a), (c) bandpass (BP) filter on the modulation spectrum, (d) modulation spectrum of BP-filtered noise, (e) amplitude envelope of (d), and (f) normalized amplitude envelope of (e).

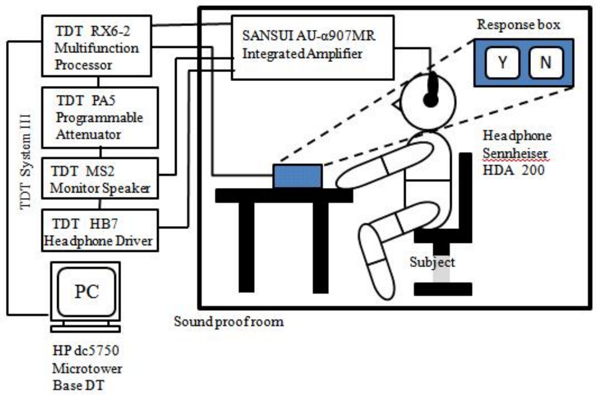

Figure 3: Environment for auditory search experiment.

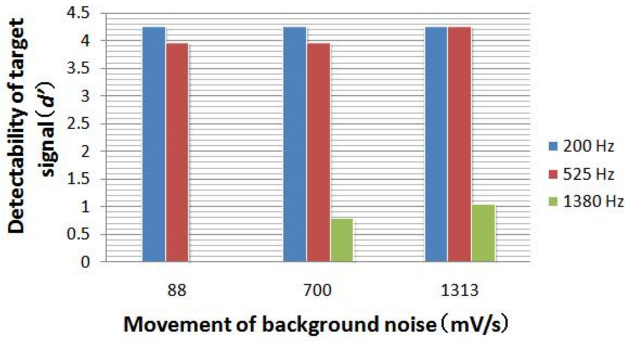

Figure 4: Detectability of target signal by the difference between movements in the temporal envelopes of target and background signals. Movement of target signal is 88 mV/s, the center frequency of target signal is 1380 Hz, and the center frequencies of background signal are 200, 525, and 1380 Hz.

pants via headphone. In the negative judgments, background signal was separately presented to the participants. The participants were required to judge whether there was the target signal with the background signal and to press "true" or "false" button on the response box. Before the participant started to run the session in this experiment, a target signal was repetitively presented to the participant to learn the target signal that would be presented in the session. All trials in the session were 720 (= nine target signals × eight background signals × two judgments (positive and negative judgments) × five repetitions). These trials were randomly presented.

4. Results

We obtained the detectability of target signal, i.e., d', from the results of the judgments of the participants. The classifications of the results of the judgments were the four types: Hit (they pressed "true" when there was a target), Miss (they pressed "false" though there was a target), False Alarm (they pressed "true" though there was not a target), Correct Rejection ((they pressed "false" when there was not a target). The greater value of d' indicates that the target signal is easily detected.

Figure 4 shows d' as a function of the movements of the background signals in which the CF and the movement of the target signal were 1380 Hz and 88 mV/s, respectively.

When the CF of target signal differed from the background signal (see blue and red bars in Fig. 4), the value of d' was constantly great regardless of the movement of the background signal. These results indicate that the target signal is easily detected. In the current experiment, three kinds of the CF of the target signal and background signals were set to be 200, 525, and 1380 Hz. Under such conditions, the target and background signals that have different CFs are resolved in the auditory periphery systems. Thus, the signal detection is very easy and the effect of the movement to detect the target signal is very small.

On the other hand, when the CF of target signal was the same as that of the background signal (see green bar in Fig. 4), the value of d' varied with the movement of the background signal, whereas the value of d' was small. Under this

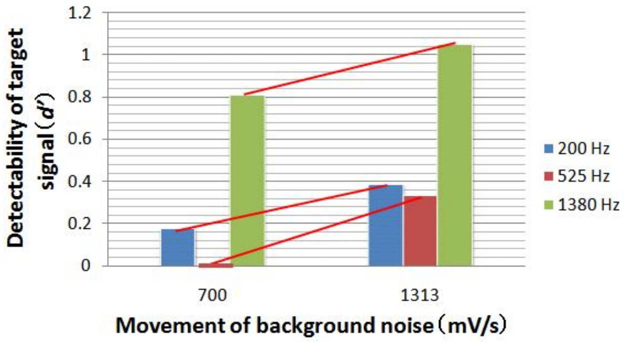

Figure 5: Detectability (d') of target signal by the difference of the movements. Movement of target is 88 mV/s and the center frequency of target is the same as that of background.

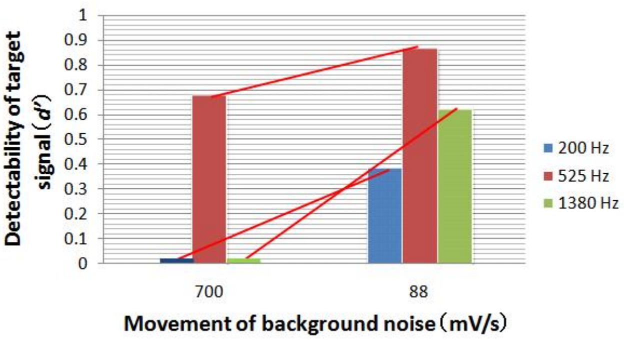

Figure 6: Detectability (d') of target signal by the difference of the movements. Movement of target is 1313 mV/s and the center frequency of target is the same as that of background.

condition, the target and background signals pass through the same auditory filter. This indicates that these two signals are decomposed into one component as one sound. Thus, it is basically difficult to separate two signals. Our study aims to improve the detectability of the target signal in such difficult situation by controlling the movement in temporal envelopes of the target and background signals. The results showed that the value of d' varied with the movement of the background signal. It suggests that the detectability of the target signal might be improved by the relationship between the movements of the target and background signals. To discussing this tendency in detail, we analyzed the experiment data in which the CFs of the target and background signals were the same.

Figure 5 shows d' as a function of the movement of the background signal when the movement of the target signal was 88 mV/s. Figure 6 shows d' as a function of the movement of the background signal when the movement of the target signal was 1313 mV/s.

These results show that the detectability (d') of the target signal is improved as the difference of the movements between the target and background signals increases on any of three kinds of CF of target signals. This tendency constantly appeared when the movement of the target signal was faster than that of the background signal. This means that we can perceptually separate the target and background signals and detect the target signals by increasing the difference of the movements between the target and background signals. Therefore, the results of the auditory search experiment suggest that the target signal is easily detected as the difference of the movements between the target and background signals increases when the CF of target signal was the same as that of the background signal.

5. Conclusion

We investigated what movements in temporal envelopes of the target and background signals can improve the detectability of the target against the background signals. We found that the detectability of target signal is improved as the difference of the movements increases in which the CF of target signal was the same as that of the background signal. This tendency constantly appeared when the movement of the target signal was faster than that of the background signal. These results suggest that greater difference of the movements plays a role of improving the detectability of the target in auditory search tasks.

Acknowledgments

This work was partially supported by Ono Acoustics Research Fund.

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," *J. 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. Sadralodabei, V. Bull, and D. R. Perrot, "Free field release from masking," *J. Acoust. Soc. Am.*, **90**(3), 1355–1370, 1991.

[5] E. G. Cherry, "Some experiments on the recognition of speech with one and with two ears," *J. Acoust. Soc. Am.*, **25**(5), 975–979, 1953.

[6] N. Asemi, Y. Sugita, and Y. Suzuki, "Auditory search asymmetry between normal Japanese speech sounds and time-reversed speech sounds distributed on the frontal-horizontal plane," *Acoust. Sci. &Tech.* **24**, 145–147, 3 (2003).

[7] M. Kusaba, M. Unoki, and M. Akagi, "A Study on detectability of target signal in background noise by utilizing similarity of temporal envelopes in auditory search," *Proc. NCSP'08*, 13–16, 2008.