| | |
|---|---|
| Title | A binaural model accounting for spatial masking release |
| Author(s) | Mizukawa, Shinya; Akagi, Masato |
| Citation | 2011 International Workshop on Nonlinear Circuits, Communication and Signal Processing (NCSP'11): 179-182 |
| Issue Date | 2011-03-02 |
| Type | Conference Paper |
| Text version | publisher |
| URL | http://hdl.handle.net/10119/9972 |
| Rights | This material is posted here with permission of the Research Institute of Signal Processing Japan. Shinya Mizukawa and Masato Akagi, 2011 International Workshop on Nonlinear Circuits, Communication and Signal Processing (NCSP'11), 2011, pp.179-182. |
| Description | |

2011 International Workshop on Nonlinear Circuits,
Communication and Signal Processing
NCSP'11, Tianjin, China, March 1-3, 2011

# A binaural model accounting for spatial masking release

Shinya Mizukawa[†] and Masato Akagi[†]

†Japan Advanced Institute of Science and Technology
1-1 Asahidai , Nomi , Ishikawa , Japan
Phone/FAX:+81-761-51-1391/+81-761-51-1390
Email: s0810060, Akagi@jaist.ac.jp

## Abstract

It is well known that detectability of signals could be improved by using directional information. This phenomenon is referred as to spatial release from masking (SRM). Interaural time difference (ITD) and interaural level difference (ILD) are used as significant spatial cues in SRM. However, the phenomenon in which ILD yields the great SRM has not yet been considered to construct models for SRM. This paper proposes a binaural hearing model that accounts for the SRM using the ITD and ILD cues. Firstly, our model divides target signals and maskers incoming two ears using a gammatone filter bank as auditory filter. Next, we consider to implement ITD and ILD processes independently and in parallel. Finally, we connect the outputs of two processes for the final masking release occurred by ITD and ILD. The results of the model show the same tendency observed as the empirical data. Consequently, our binaural hearing model can simulate the SRM.

## 1. Introduction

There are cases that target signal cannot be perceived in real environments because target signals are possibly masked by varieties of noises. For signal detection in the noisy environments, Ebata *et al.* reported that detectability of signals could be improved using directional information [1]. In addition, Saberi *et al.* found that detectability of signals was improved in a free sound field when the signal and masker were spatially separated [2]. ITD and ILD are used as significant spatial cues in SRM. On the other hand, according to the frequency components of signal, binaural masking level difference (BMLD) occurs at the same time as SRM. Interaural phase difference (IPD) is used as a significant cues in BMLD. Kuroda *et al.* carried out experiments using signals with car noise which has ITD , IPD and ILD [3]. The results revealed when a target signal is low in frequency (about 1.0 kHz), detectability of the target signal is improved by ITD and IPD. Moreover, ILD yields great SRM when the target signal is high (about 2.5 kHz) and is fixed in front of listeners. These phenomena play an important role to detect alarm signals. However, how these phenomena are occurred by ITD, IPD and ILD in auditory system is still an open question. The purpose of our study is to investigate a binaural hearing model accounting for masking release by the significant spatial cues and to clarify how SRM occurs using the ITD, IPD and ILD.

## 2. Equalization - Cancellation Model

Durlack proposed the equalization-cancellation (EC) model to interpret on BMLD [4]. The EC model assumed that the auditory system transforms a masker incoming one ear into that incoming other ear. Then, the equalized maskers cancel each other using ITD (IPD) cues to eliminate the masker components. Signal-to-masker ratio with the cancelled maskers increases and then detectability of the target signals goes up. The facter can be expressed as

$$f_m = \frac{(1+a^2)(1+\sigma_\epsilon{}^2) - 2a\cos(\omega\tau)\exp(-\omega^2\sigma_\delta{}^2)}{2[1+\sigma_\epsilon{}^2 - \exp(-\omega^2\sigma_\delta{}^2)]} \quad (1)$$

The output $f$ is called EC facter. The EC facter shows the signal-to-masker ratio after EC process. The EC model calculates BMLD using the EC facters.

$$\mathrm{BMLD} = 10\log_{10}\frac{f_m}{f_0} \quad (2)$$

where, $f_0$ is the EC facter when the target siganl and the masker are fixed at the front, referred as $S_0N_0$. $f_m$ is the EC facter when the target signal or the masker are fixed at the front, $S_0N_m$ or $S_mN_0$.

We simulate masking releases occurred by ITD with the same experimental signals and maskers as used by Nakanishi *et al.* [5]. ITD is calculated by

$$\mathrm{ITD} \quad = \quad \frac{d}{c} = \frac{r(\theta+\sin\theta)}{c} \quad (3)$$

where $r$ in meters is the radius of the head, $\theta$ in radians is the direction sound source, $c$ in meters per secound is the sound velocity and $d$ in meters is the path difference from the sound source to both ears. In this simulation, $r$ is set to 0.09 m and $c$ is set to 343.5 m/s. Directions of the target signal are varied from 0 ° to 90 ° in step of 1 ° , and the direction in front of

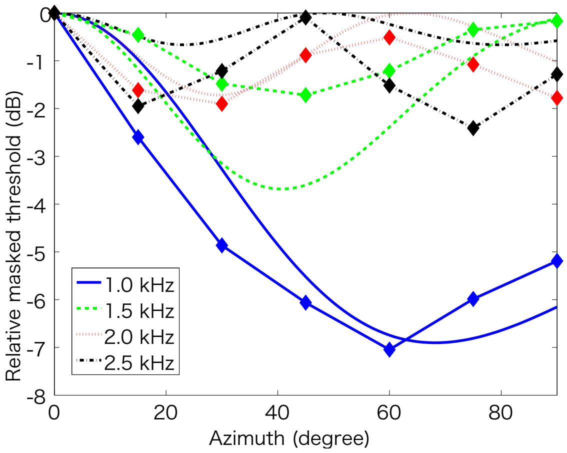

Figure 1: Lines with no markers are outputs of the EC model, lines with markers are empirical data by Nakanishi *et al.* [5].

listners is 0°. Frequency of the input signals are varied from 1.0 kHz to 2.5 kHz in step of 0.5 kHz. The error variances of the time and amplitude $\sigma_\epsilon$ and $\sigma_\delta$ for the EC model are set to $\sigma_\epsilon = 0.25$ and $\sigma_\delta = 105 \times 10^{-6}$.

Figure 1 plots the results for the mean masked thresholds in each direction. Lines with no markers are outputs of the EC model, lines with markers are empirical data by Nakanishi *et al.* [5]. The results described similar characteristics of the masking release occurred by ITD. The EC model could estimate the masking release occurred by ITD with high accuracy. However, when the input signals have ILD greatly, the accuracy of the simulations goes down. The phenomenon in which ILD yields the great SRM has not yet been considered to implement the model.

## 3. implementation of the model

In this paper, we propose a binaural hearing model that accounts for the spatial masking release using the ITD and ILD cues. The model is based on EC model. Figure 2 shows the flow of the proposed model.

In human sound localization mechanism, ITD and ILD are detected by alternative organs. Then, we consider to implement ITD process and ILD process independently and in parallel. Input of the model is the binaural signal and the masker shown as $S_l$, $S_r$, $N_l$ and $N_r$. Output of the model is the masking release occurred by ITD and ILD.

### 3.1. Frequency division

Firstly, we discuss to compute amount of masking release occurred complex tone hearing. Complex tones have different ILDs on bands of frequency. Thus SRM is different among frequency bands. Our model divides the target signals and maskers incoming two ears into the frequency bands using a filter bank. The model adopts a gammatone filter bank as auditory filter. Figure 3 shows amplitude gains of the gammatone filter bank that dividing into 20 channels whose minimum center frequency is 0.5 kHz and maximum center frequency is 3.0 kHz. Then, SNR is calculated for each band. Bands having low SNR are neglected.

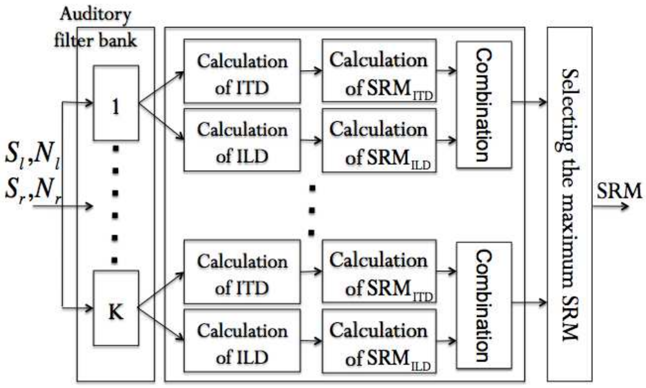

Figure 2: The flow of the proposed model

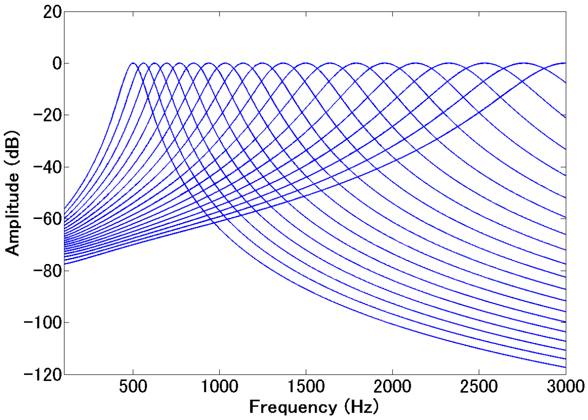

Figure 3: Gammatone filter bank

### 3.2. Calculation of ITD and ILD

The output of the filters are used to calculate ITDs and ILDs for each band. ITD is calculated using cross correlation between left and right signals, using the following equation,

$$R_{lr}(\tau) = \int_{-\infty}^{\infty} S_l(t) S_r(t-\tau) dt \tag{4}$$

where $R_{lr}(\tau)$ is the maximum value, when $\tau$ indicates ITD.

ILD is calculated by power ratio between left and right signals.

$$a = \frac{\int_0^T S_r^2(t)dt}{\int_0^T S_l^2(t)dt} \tag{5}$$

ILD is referred as $a$ [dB].

### 3.3. Masking release occurred by ITD and ILD

In this section, we propose a method for calculating amounts of masking release occurred by ITDs and ILDs for each band. The calculating method is based on EC theory. We assume the equalization occurred by ITD and the enhancement by ILD in the signal detection mechanism.

Firstly, we consider to equalize the two maskers using ITD and ILD values. If differences of the two maskers are only ITD and ILD, the two maskers are formulated as the following

$$N_l = aN_r(t-\tau) \tag{6}$$

where $\tau$ is ITD, $a$ is ILD.

In ITD process, the equalized maskers are cancelled each other using ITD cues to eliminate the maskers. To equalize the two maskers, delay $\tau$ is added one masker. The signal on the same side is added to the same delay. Then, the masker in one ear is cancelled by subtracting the masker in other ear. However, the signals have different added delay to equalize the two maskers, the target signal components can't be cancelled completely. After the cancellation, signal-to-noise ratio $f_{\tau m}$ indicates detectability of the signal. $f_{\tau m}$ is computed by

$$f_{\tau m} = \frac{\int_0^T (S_l(t) - S_r(t-\tau-\sigma_\delta))^2}{\int_0^T (N_l(t) - (a+\sigma_\epsilon)N_r(t-\tau-\sigma_\delta))^2} \tag{7}$$

where $\sigma_\epsilon$ and $\sigma_\delta$ are error variances of the time and amplitude. SRM is calculated to compare $f_{\tau 0}$ in $S_0N_0$ with $f_{\tau m}$ in $N_0S_m$.

$$\text{SRM}_\tau = 10\log_{10}\frac{f_{\tau m}}{f_{\tau 0}} \tag{8}$$

$\text{SRM}_\tau$ is the masking release occurred by using ITD.

In ILD process, target signals are enhanced using ILD cues to rise up to exceed the masker. To equalize the two maskers, enhancing power$a$ is multiplied to one masker. The signals are enhanced to be the same power. Then, the masker in one ear is cancelled by subtracting the masker in other ear. Because the signals is enhanced power to equalize the two maskers, the target signal components can't be cancelled completely. The signal-to-noise ratio $f_{am}$ indicates detectability of the signal, and $f_{am}$ is computed by

$$f_{am} = \frac{\int_0^T (S_l(t) - (a+\sigma_\epsilon)S_r(t))^2}{\int_0^T (N_l(t) - (a+\sigma_\epsilon)N_r(t-\tau-\sigma_\delta))^2} \tag{9}$$

SRM is calculated to compare $f_{a0}$ in $S_0N_0$ with $f_{am}$ in $N_0S_m$ by the same ITD process.

$$\text{SRM}_a = 10\log_{10}\frac{f_{am}}{f_{a0}} \tag{10}$$

$\text{SRM}_a$ is the masking release occurred by ILD. The two processes can compute amount of masking release occurred by ITD and ILD separately. The model can deals with the masking release from not only the effectiveness of ITDs but also the effectiveness of ILDs.

### 3.4. Combination the two processes

We consider to combine the outputs of two processes for the masking release occurred by ITD and ILD. As explained in the previous section, we defined ITD and ILD processes independently in our model. Therefore, the model needs a weight to mix effects of ITD and ILD at the frequency of the signal. It is well known that ITD effects availability of sound localization in low-frequency (below 1.5 kHz) signal and ILD effects availability of sound localization in high-frequency (above 2.0 kHz) signal. Here, the model refers the relation between the characteristic frequency (CF) and synchronization index of auditory nerve in the acoustic system. In the acoustic system, auditory nerve firing pattern is synchronized into phase of the stimulus. The synchronization index is high with the low-frequency stimulus, and low with the high-frequency stimulus. Jhonson [6] reported that synchronization index is about 0.7 to 0.9 with CF below 1.0 kHz. Synchronization index is droped to a lower value with an rise in CF, and availability of sound localization using ITD goes down. We assume that the effectiveness of ITD and the effectiveness of ILD is complementary. As effectiveness of ITD decreases, the effectiveness of ILD increases. The model adopts a fitted curve by described Blackburn and Sachs [7] by following formula for the weight to mix effects of ITD and ILD,

$$\alpha = 0.97 - 0.16cf - 0.01cf^2 \tag{11}$$

The total SRM is determined by weighting $\text{SRM}_\tau$ calculated using ITD and $\text{SRM}_a$ calculated using ILD.

$$\text{SRM} = \alpha\text{SRM}_\tau + (1-\alpha)\text{SRM}_a \tag{12}$$

SRM is calculated for each band. The final output of the model is the largest masking release selected from all outputs of the subbands.

## 4. Simulation

In this section, we simulate the SRM and BMLD to evaluate the model.

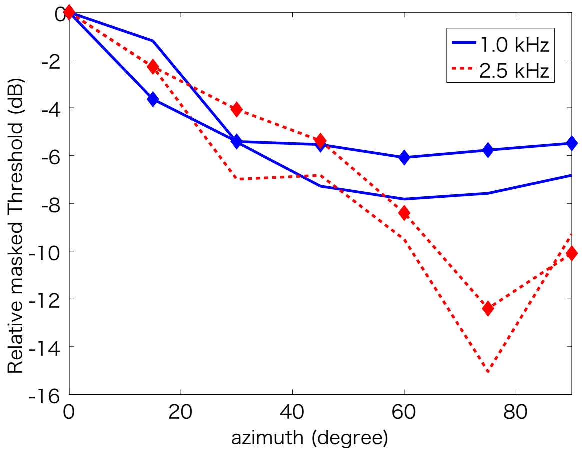

Figure 4: The result of (a), lines with no markers are outputs of the model, lines with markers are empirical data by Kuroda *et al.* [3].

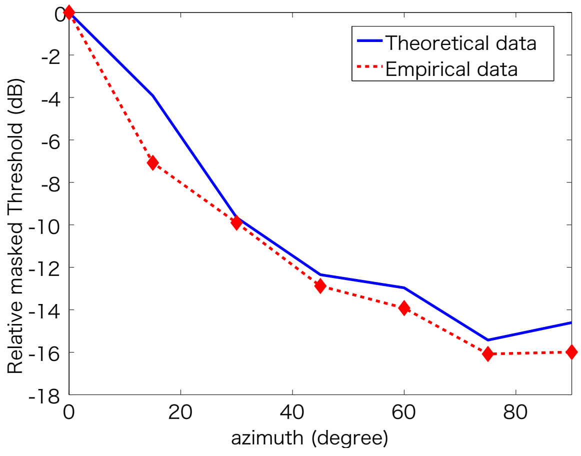

Figure 5: The result of (b), lines with no markers is output of the model, lines with markers is empirical data by Kuroda *et al.* [3].

### 4.1. Configuration of simulation

This simulation data is the same signals and maskers as used by Kuroda *et al.* We follow the following two experiments as the simulation.
(a) Target signal is puretone whose frequency component is 1.0 kHz and 2.5 kHz, and masker is car noise.
(b) Target signal is a pulse train signal, and masker is car noise.
Directions of presentation of the masker are varied from 0 ° to 90 ° in step of 15 ° . The error variances $\sigma_\epsilon$ and $\sigma_\delta$ in the model are set to $\sigma_\epsilon = 0.25$ and $\sigma_\delta = 105 \times 10^{-6}$.

### 4.2. Results and discussions

Figure 1 shows the result of the condition (a). Figure 2 shows the result of the condition (b). The vertical axis indicates the relative masked thresholds, that have been normalized by the masked thresholds at $S_0N_0$. The horizontal axis indicates the azimuth. The results of the simulation using the proposed model showed the same tendency as the empirical data by Kuroda *et al.* [3]. Consequently, our binaural hearing model could calculate the masking release occurred by ITD and ILD.

## 5. Conclusions

In this paper, we proposed a binaural hearing model that accounts for the spatial masking release. The model is implemented to base on assumption occurred the elimination of the maskers in the acoustic system. The results of the simulation show similar amount of the masking release to the empirical data. This finfing suggest that our binaural hearing model can calculate the masking release occurred by ITD and ILD.

## References

[1] M. Ebata, T. Sone, and T. Nimura, “Improvement of hearing ability by directional information,” J. Acoust. Soc. Am, Vol. 43, pp. 289-297, 1968.

[2] K. Saberi, L. Dostal, T. Sadralodabai, V. Bull, D.R. Perrott “Free-field release from masking,” J. Acoust. Soc. Am, Vol. 90, no. 3, pp. 1355-1370, 1991.

[3] N. Kuroda, L. Junfeng, Y. Iwaya, M. Unoki, M. Akagi “Effects from spatial cues on detectability of alarm signals in noisy environments,” IWPASH, pp. 1-4, Nov.11-13, 2009

[4] N. I. Durlach, “Equalization and cancellation theory of binaural masking level differences,” J. Acoust. Soc. Am, Vol. 35, pp. 1206-1218, 1963.

[5] J. Nakanishi, M. Unoki, M. Akagi “Effect of ITD and Component Frequencies on Perception of Alarm Signals in Noisy Environments, ” Jounal of Signal Processing, Vol.10-4, pp.231-234, 2006

[6] Don H. Johnson, “The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones” J. Acoust. Soc. Am. Vol. 68, Issue 4, pp. 1115-1122, 1980

[7] C. C. Blackburn and M. B. Sachs, “Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis” Journal of Neurophysiology, Vol 62, Issue 6 1303-1329