A COMPUTATIONAL MODEL OF AUDITORY SOUND LOCALIZATION BASED ON ITD

KAZUHITO ITO AND MASATO AKAGI

Japan Advanced Institute of Science and Technology, Hokuriku
1-1 Asahidai, Tastunokuchi, Nomigun, Ishikawa 923-1292 Japan
E-mail: keith@jaist.ac.jp akagi@jaist.ac.jp

This paper introduces a functional model of auditory sound localization based on the interaural time difference (ITD). The signals in the nervous system such as action potentials and synaptic transmission are modeled computationally and these applied to a coincidence detector circuit model to detect ITD. Then impulse trains fluctuating in time are used as input. The incorporated model outputs a spike histogram of which the peak of the envelope will indicate the ITD. The simulation results show that a peak indication the ITD in the azimuth clearly sharpens when using impulses fluctuating in time as input. This suggests that impulse fluctuation does not behave like noise and it can contribute to the detection of ITDs in the temporally redundant process and the nonlinear output mechanism.

1 Introduction

Sound localization based on the interaural time difference (ITD) is an auditory function for detecting sound source locations using the temporal difference between the ears at which the sound waves arrive [1][2][3]. To understand the process and to represent it computationally, we developed a computational model of the auditory sound localization based on the ITD. We have focused in particular on the temporal redundancy of signals in the nervous system because it may affect temporal information for ITDs. We also looked at the process of detecting ITDs when using impulse trains with fluctuation in time as input data to the model. The impulse fluctuation has tended to be regarded as noise in the traditional logic of the binaural system. We suggest that it may be useful.

2 Cross Correlation

In general, traditional models for detecting ITDs, like the Jeffress model [4], have been represented as a circuit consisting of some coincidence detectors and two nerve fibers from the left and right ears. The detectors fire the most when impulses from both sides arrive at the detector simultaneously. Such a model can calculate ITDs by cross-correlating the impulse trains from both sides [5]. Structures similar to those of the coincidence detector circuit model are found in the real auditory system. For example, as the medial superior olivary and the nucleus laminaris [2][3]. Impulses inputted into these structures originate in the auditory nerve fibers

(ANFs) through the bushy cells whose responses are called primary-like. The impulse trains have temporal information, because the impulses from ANFs are in synchronization with a certain phase of the stimuli. Therefore, the cross-correlation model is considered as a basic system for detecting ITDs.

3 Impulse Fluctuation

However, ANFs do not always fire in synchronization with a certain phase of stimuli and the impulses fluctuate slightly in time [6]. If we use impulse trains with a fluctuation in time from the auditory peripheral model [7], the coincidence detector circuit model will be affected by the impulse fluctuation. The fluctuation will behave like noise in the cross-correlation model. Impulse trains having a characteristic frequency of 300 Hz with a large fluctuation in time and duration of 0.3 s are provided and input to the cross-correlation model with a time difference of 100 μs. The upper panel in Figure 1 shows the period histogram of one of those impulse trains with a large fluctuation and the lower panel shows the spike histogram obtained by this simulation. The spike histogram has some peaks but the peaks do not indicate the ITD. Impulse fluctuation in input data makes it difficult for the cross-correlation model to determine the ITD. The impulse fluctuation has therefore been regarded as noise. To solve this problem of the effect of noise, one approach is to set many coincidence detector circuits and integrate all the outputs to improve the robustness of detecting ITDs.

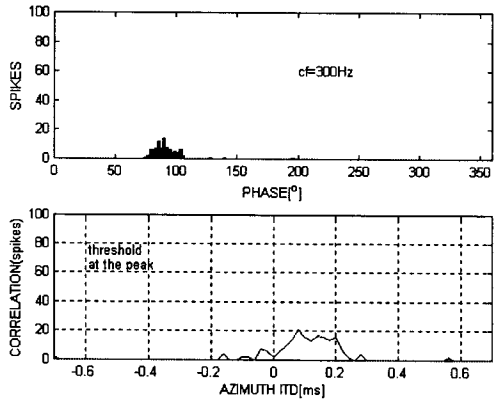


Figure 1. Period histogram of the impulse train with a large fluctuation in time and the spike histogram as obtained by cross correlation (ITD =100 μs). The envelope of the spike histogram has some peaks.

4 Temporal redundancy in the signals

We focus on signaling in the real auditory system. Since structures similar to those of the coincidence detector circuit are found in the system, it is natural that the signals in the nervous system, such as synaptic transmission, are used in the

circuit. Additionally, the duration of these signals are much longer than the minute ITD perceived by humans. For example, humans can perceive an ITD variation of about 10 μs at 900 Hz, corresponding to a minimum audible angle (MAA) of about 1° [1][8]. The duration of synaptic transmission ranges from several milliseconds to hundreds of milliseconds.

Figure 2 shows two types of temporal transition of postsynaptic potentials (PSPs) on two coincidence detectors corresponding to the detection of 0-μs and 20-μs ITDs. An impulse from each side is input to the circuit without any time difference. For a human to perceive an ITD variation of about 10 μs as the MAA, the duration of the PSP should be shorter than 10 μs (Fig. 2A). Since impulses arrive at the 0-μs ITD detector at the same time, two PSPs combine together and make a large potential. Since the arrival times of impulses at the 20-μs ITD detector do not match, PSPs decline without affecting each other. Thus it is easy to distinguish the difference between the two detectors and determine ITDs. However, it should be considered that the duration of a PSP by synaptic transmission is several milliseconds or more (Fig. 2B). At the 0-μs ITD detector, two long PSPs combine together and make a large potential. Likewise, at the 20-μs ITD detector, two long PSPs combine together and make a large potential, even though the arrival times of the two impulses do not match, because the temporal interval between them is much smaller than the duration of PSPs. Thus, the duration of PSPs results in temporal redundancy and it would obscure minute information such as the temporal interval between the arrival times of impulses.

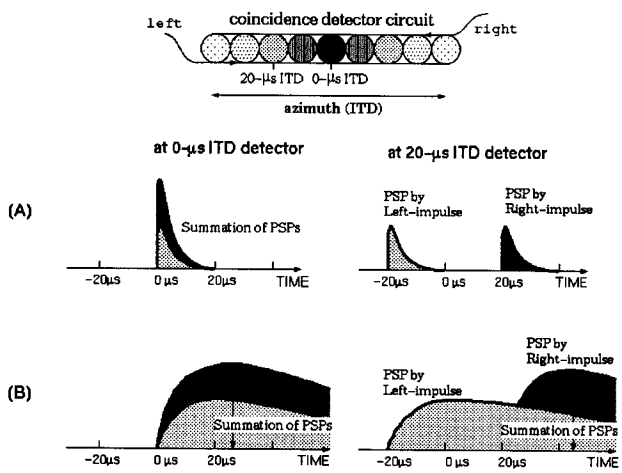


Figure 2. Temporal redundancy.

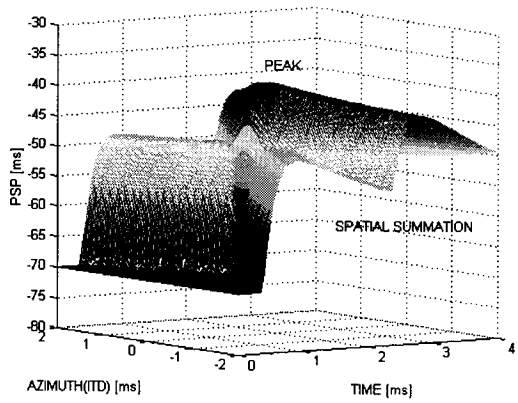


Figure 3. Postsynaptic potentials in a coincidence detector circuit.

In this study, the signals are modeled computationally using Hodgkin-Huxley equations [7][9] to represent the temporal redundancy. Although applying these equations might not be sufficient to describe the complicated process in the nervous system, this is one way to represent the temporal redundancy of signals. Synaptic transmission is represented by the following equations.

$$C_m \frac{dV(t)}{dt} = -g_{\mathrm{Na}}(V(t) - E_{\mathrm{Na}}) - g_{\mathrm{K}}(V(t) - E_{\mathrm{K}}) - g_L(V(t) - E_m) \ , \qquad (1)$$

$$g_n(t) = A_n(t - t_n)e^{-(t-t_n)/\tau_n} \ , \qquad (2)$$

where $V(t)$ is the postsynaptic potential (PSP) at time t, C_m is the membrane capacitance, E_m is the resting potential, E_{Na} and E_{K} are the equilibrium potentials for

sodium and potassium, g_{Na}, g_{K}, and g_L are the postsynaptic conductances for sodium, potassium, and the leakage, respectively. $g_n(t)$ indicates the postsynaptic conductance of ion n at time t. t_n is the time of the most recent onset of ion n conductance. τ_n is the time constant for the conductance, and A_n is the amplitude constant related to the permeability of ion n.

In this model, the function for firing is separated from the potential function. To represent the effects of other conductances, such as early potassium channels, voltage-gated calcium channels, and calcium-activated potassium channels, the firing threshold level is varied according to Eq. (3), which translates the magnitude of the grand PSPs into the firing frequency of action potentials

$$V_{\mathrm{threshold}}(t) = \beta \cdot e^{-(t-t_r)/\tau_r} + E_{\mathrm{threshold}} \ , \qquad (3)$$

where $V_{\mathrm{threshold}}(t)$ indicates the threshold level at time t and t_r is the time of the most recent discharge. $E_{\mathrm{threshold}}$ is the basis of the threshold level in this function, τ_r is the time constant representing the relative refractory period and the adaptation, and β is the maximum amplitude constant. These modeled signals were thus applied to a coincidence detector circuit to detect ITDs. Figure 3 shows the temporal transition of PSPs on the coincidence detector cells arranged in the azimuth (the axis of ITDs). One impulse from each side was input to the circuit without any time difference. The impulses from each ear stimulated the coincidence detectors sequentially from each side of the circuit and a small PSP was generated on every detector. When impulses from both sides met at the middle detector in azimuth simultaneously, the PSPs were summed and a large potential was generated. Then the impulses separated and kept stimulating other detectors on the other side, and summations of PSPs were generated on both sides. The envelope with the maximum potential on every detector draws a peak on the axis of ITDs. The peak should indicate the detector of the ITD but it is too broad to determine it. Since the neighboring detectors also had large potentials close to the peak potential, there was not a clear difference to uniquely distinguish the detector of the ITD from the others.

5 Nonlinear Output Mechanism

In this model, the firing threshold function (Eq. 3) is established to represent a neuron firing. It is difficult to satisfactorily set the threshold level to the same level

as the peak of the potential envelope. It is natural to set it to a level below the peak. In that case, all the detectors whose potential exceeds the threshold level fire and a certain range of firing appears on the azimuth (Figure 4). Accordingly, the model represents a nonlinear output mechanism. The model outputs spikes over a broad range along the axis of ITD because of the broad peak of the PSP. This is another temporal redundancy. Figure 5 shows the result of a simulation by the nonlinear output mechanism using the impulse trains firing in synchronization with a certain phase of stimuli. The upper panel in Fig. 5 shows the period histogram of the impulse train and the lower panel shows the spike histogram obtained by this simulation. When impulse trains firing in synchronization with a certain phase of stimuli were used as input data, the model kept outputting spikes over the fixed broad range, and the output spikes were distributed widely and evenly on the axis of ITD. This model does not appear to uniquely determine the actual ITD. The result suggests that representing the temporal redundancy and the nonlinear output mechanism is not sufficient for the detection of ITDs. We have two drawbacks here: One is the impulse fluctuation that affects the detection of ITDs in the cross-correlation model. The other is the temporal redundancy of the signals that makes outputs from a coincidence detector circuit broad.

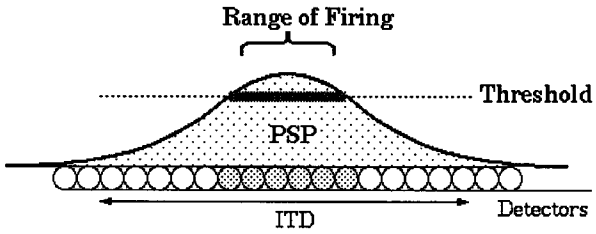
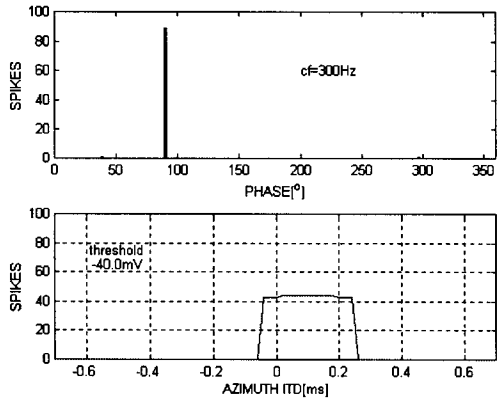


Figure 4. Threshold level below the peak of the potential envelope.



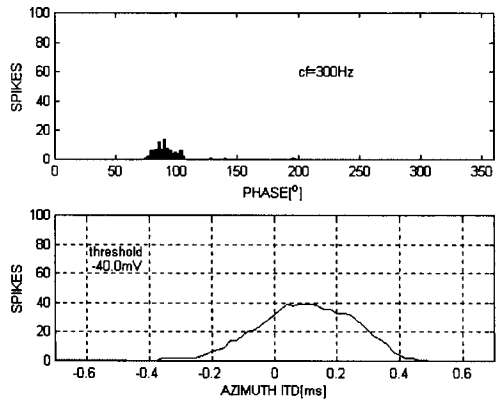


Figure 5. (left) Period histogram of the impulse train firing in synchronization with a certain phase of stimuli and the spike histogram obtained by the nonlinear output mechanism (ITD = 100 μs). The envelope of the spike histogram can not indicate the ITD.

Figure 6. (right) Period histogram of the impulse train with a large fluctuation and the spike histogram obtained by nonlinear output mechanism (ITD =100 μs). The envelope of the spike histogram on the ITD rises.

6 Overcoming the two drawbacks

Each of these drawbacks by itself makes it impossible to detect the ITD. However, we will show that combining these drawbacks produces a good result. We show the result of the simulation when impulse trains with a large fluctuation in time are used as input data. Figure 6 shows the result of a simulation by the nonlinear output mechanism using the same impulse trains as those in Fig. 1. The upper panel in Fig. 6 shows the period histogram of the impulse train with a large fluctuation and the lower panel shows the spike histogram obtained by this simulation. The envelope of the spike histogram on the ITD is rising and beginning to form a peak.

We describe the principle. Impulse trains with a large fluctuation in time are used in the model with the nonlinear output mechanism by setting the firing threshold level below the peak of the potential envelope. Impulse fluctuation in time affects the location where the broad peak of the PSP envelope appears on the axis of ITD, because the detector where impulses from both sides of the circuit encounter each other changes whenever impulses are input to the model. The location of the peak of the PSP envelope varies and the firing range also shifts along the axis of ITD (Figure 7A). If the threshold level is set to an appropriate level, there is an area where the firing ranges overlap each other, in spite of the variation in the peak location. The detectors in the overlapping area tend to keep firing and output more spikes. Accordingly, the output spikes are distributed like the Gaussian distribution on the axis of ITD. The distribution has a peak indicating the actual ITD uniquely.

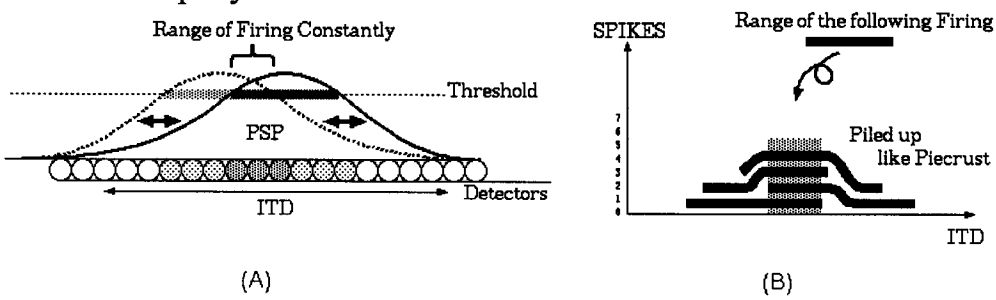


Figure 7. Nonlinear output mechanism.

If a spike histogram is drawn according to the variation in the firing range, the number of spikes in the overlapping area is found to increase like piled-up pie crust faster than in other areas (Figure 7B). This means that using impulse trains with fluctuation in time as input data improves the output of the model compared with using impulse trains that have no fluctuation, and the envelopes of spike histograms output from the nonlinear output mechanism tend to have a peak that indicates the ITD or its vicinity. The distinct difference between the traditional model and the presented model is the way of thinking about impulse fluctuation in input data. The traditional model regards impulse fluctuation as noise, while the presented model regards it as something useful in the temporally redundant process

and the nonlinear output mechanism. These concepts oppose one another.

7 Conclusion

A computational model of the auditory sound localization based on the interaural time difference (ITD) was presented. The signals in the nervous system were modeled computationally and these models were applied to a coincidence detector circuit model to detect ITDs. Impulse trains with fluctuation in time were used as input data and the effects of the impulse fluctuation on the detection of ITDs were investigated. The simulations show that using impulse trains with fluctuation in time as input data improves the outputs of the model compared with using impulse trains having no fluctuation. This suggests that impulse fluctuation does not behave like noise and that it can contribute to the detection of ITDs in the temporally redundant process and the nonlinear output mechanism.

8 Acknowledgements

This work was supported by CREST and by Grant-in-Aid for Science Research from Ministry of Education (No.10680374).

References

[1] Moor, B. C. J. (1997). *An Introduction to the Psychology of Hearing.* Academic Press, Cambridge.

[2] Konishi, M. (1993). "Listening with Two Ears." *Scientific American*, pp. 34-41.

[3] Kuwada, S., Batra, R., and Fitzpatrick, D. C. (1997)."Neural Processing of Binaural Temporal Cues," in *Binaural and Spatial Hearing in Real and Virtual Environments*, R. H. Gilkey and T. R. Anderson, Lawrence Erlbaum Associates, New Jersey, pp. 399-425.

[4] Jeffress, L. A. (1948). "A Place Theory of Sound Localization." *J. Comp. Physiol. Pychol.*, **45**, pp. 35-49.

[5] Stern, R. M. and Trahiotis, C. (1995). "Models of Binaural Interaction," in *Hearing*, B.C.J. Moor, Academic Press: California, pp. 347-386.

[6] Johnson, D. H. (1980)."The relationship between spike rate and synchrony in responses of auditory nerve fibers to single tones." *J. Acoust. Soc. Am.* **68**, pp. 1115-1122.

[7] Maki, K. and Akagi, M. (1997). "A Functional Model of the Auditory Peripheral System." *Proc. ASVA97*, Tokyo, pp. 703-710.

[8] Mills, A.W. (1958)."On the minimum audible angle." *J.Acoust.Soc.Am.*, vol. **30**, pp. 237-246.

[9] Rothman, J. S., Young, E. D., and Manis, P. B. (1993). "Convergence of Auditory Nerve Fibers Onto Bushy Cells in the Ventral Cochlear Nucleus: Implications of a Computational Model." *J.Neurophysiolgy*, vol. **70**(6), pp. 2562-2583.