Echolocation of Multiple Targets in 3-D Space from a Single Emission

M. YANO¹, I. MATSUO¹ and J. TAN²

¹*Research Institute of Electrical Communication, Tohoku University, Japan*

²*Institute of Fluid Science, Tohoku University, Katahira 2-1-1, Aoba-ku, Sendai, 980-8577, Japan*

Abstract. Using frequency-modulated echolocation sound, bat can capture a moving target in real three-dimensional (3-D) space. It is impossible to locate multiple targets in 3-D space by using only the delay time between an emission and the resulting echoes received at two points (i.e., two ears). To locate multiple targets in 3-D space requires directional information for each target. The spectrum of the echoes from nearly equidistant targets includes spectral components of both the interference between the echoes and the interference resulting from the physical process of reception at the external ear. The frequency of the spectral notch, which is the frequency corresponding to the minimum of the external ear's transfer function (EEDNF), provides a crucial cue for directional localization. In the model we present, a computational model to discriminate multiple close targets in 3-D space utilizing echoes evoked by a single emission by distinguishing the interference of echoes from each object and the EEDNF corresponding to each target.

Key words: acoustic image, brain, computational model, delay-tuned neuron, frequency modulated sound, echolocation, recognition

1. Introduction

Bats have the special ability to form an image of the world based on acoustic information, while most animals form images of the world that are more strongly based on visual information. They emit trains of high-frequency sounds and can locate an individual target among multiple objects by using the echoes of these emissions that overlap in time [1, 2]. In order to recognize objects in three-dimensional (3-D) space, bats need to process the information that the echoes contain to determine the distance and direction to each object.

With regard to the determination of distance, it has been found in physiological studies that delay-tuned neurons in the central auditory system are capable of determining the interval between an emitted sound and the returning echo [3–7]. In a case of interfering echoes, responses of neurons in the inferior colliculus (IC) depend on the amplitude change based on the interference between echoes [8]. There might be the mechanism that a bat uses to determine a target's distance. In fact, when two different targets are presented, either simultaneously or sequentially on either side of the animal, the accuracy of the determination of the difference in

distance to the two objects is about 1–2 cm, corresponding to a time resolution of about 50–100 μs [9–11]. When two phantom targets, one follows the bat's emissions after a fixed interval and the other follows the bat's emissions after intervals whose lengths fluctuate about some average value (a 'jitter condition'), are presented, bats can distinguish objects separated by distances of well under 1 mm [12].

The localization of targets in 3-D space has been extensively investigated in computational studies. In order to distinguish multiple targets, in particular two targets closely located on a one-dimensional range axis, Saillant et al. [13] proposed a monaural model referred to as the SCAT (spectrogram correlation and transformation) model. This SCAT model can be easily extended to a binaural model, and such an extended model could locate a single target in a two-dimensional plane. In this model, however, to locate a single target in 3-D space requires sampling of echoes from at least two different aspect angles. In the case of more than two targets, there arises the problem of the ambiguity of the correspondence between perceived echoes and objects. Theoretically, it is not yet known how such a determination of the location of multiple targets in 3-D space can be made from echoes produced by a single emission.

Determining the locations of multiple targets in 3-D space requires independently fixing the direction as well as the distance of each target. Recently, it has been found that there are spectral notches produced at each external ear, which are local minima in the amplitude of echoes [14, 15]. It is called the 'external ear dependent notch frequency (EEDNF)'. The EEDNF varies with the direction of a sound source and this is an essential cue for the determination of target direction [16,17]. In the case of multiple nearly equidistant targets, the echo spectrum includes contributions from the interference between echoes from these targets and from the interference created through the physical process of reception at the external ear [18]. In order to locate each object in 3-D space, it is necessary to separate the contributions from the interference created through the physical process of reception at the external ear and the interference between echoes from the targets. In this paper, we propose a model to discriminate multiple nearly equidistant targets in 3-D space by analyzing the echoes produced by a single emission and determining the difference between delay times and the EEDNF corresponding to each object.

2. Model

We propose a model capable of determining the positions of multiple nearly equidistant targets using only the overlapping and interfering echoes from each target resulting from a single emission. The distances from a bat to objects in a target region, including single or multiple closely positioned targets, are determined by the delay time between the emission and the return of the group of echoes. Also, the spectrogram of echoes from multiple closely positioned targets includes contributions from both the interference between the echoes from different nearly

Figure 1. Diagram of the model. The model consists of four blocks. The inputs to the cochlear block are the emission itself and the echoes reflected from targets.

equidistant targets and interference created through the physical process of reception at the external ear. The component of the spectral pattern dependent on the interference between the echoes from the targets is called the 'target dependent spectral pattern (S_{target})', and the component of the spectral pattern dependent on the interference caused by the physical process of reception at the external ear is called the 'external ear dependent spectral pattern ($S_{\text{ext_ear}}$)'. The S_{target} and the $S_{\text{ext_ear}}$ provide information for determination of the targets' distances and directions, respectively. To locate each target, it is necessary to distinguish them within the spectrogram of echoes.

As shown in Figure 1, our model consists of four processing blocks. Inputs to the cochlear block are the emission and echoes reflected from objects. The output of the cochlear block is parallel processed through the two pathways, the temporal and spectral blocks. The function of the temporal block is to determine the distance to the target region, and that of the spectral block is to determine

both the differences between delay times of echoes from nearly equidistant targets and the EEDNF associated with each target. Finally, the function of the binaural integration block is to determine a targets location in 3-D space by using cues containing information about the distance to each target and the EEDNF for each target.

EMISSION AND ECHO

In our model, we use FM emissions with a constant amplitude from 20 to 120 kHz, sweep time of 3.6-ms-long and rise/fall time of 50 μ s. We assume no directional dependence of emitted wavelengths for simplicity. We assume the following conditions to simplify the simulation. (i) Each target is a point target. (ii) The reflecting target is a perfect reflector in the sense that it does not modify the phase or amplitude of the sound wave. (iii) The emitting mouth is midway between two ears, and the distance between the mouth and each ear is 2 cm. This corresponds to the small head of an echolocating bat.

To simplify our simulation, the EEDNF (in kHz) as a function of the azimuthal and elevational angles (in degrees) of a sound incident to an ear is defined by making reference to the measured transfer function of the external ear of the bats, *Eptesicus fuscus* [15] as follows:

$$\text{EEDNF}(A, E) = CA/2 - E/5 + 40 \quad (1)$$

where A and E are the azimuthal and elevational angles of the incident sound. Since the EEDNF is assumed to possess bilateral symmetry, the value of C is 1 for the left ear and -1 for the right ear. Because azimuthal directions from -8° to 8° and elevational directions from 0° to 20° are considered in this simulation, the EEDNF is varied from 31 to 44 kHz. This frequency region is called the ‘external ear dependent notch frequency region’ (EEDNF region). The frequency region outside this region is called the ‘non-EEDNF region.’ A transfer function of the external ear is modeled by a second order infinite impulse response (IIR) filter. The filters cutoff bandwidth is 3 kHz, and the attenuation of the amplitude at the notch frequency is 20 dB with respect to the maximum.

The polar coordinate axes (r, θ, ϕ) are used to represent a target’s location in 3-D space and the mouth is fixed at the origin of the coordinate system. Then, r , θ , and ϕ are the distance, azimuthal angle, and elevational angle, respectively. Figure 2 shows an example of the waveform of emission and the returning echoes, in which three targets are located at (0.730 m, 0° , 0°), (1.000 m, 0° , 0°), and (1.040 m, 0° , 0°).

2.1. THE COCHLEAR BLOCK

The function of the cochlear block is to transform the waveforms of the emission and the echoes into the spectrograms. The processing of the bat’s inner ear can

Figure 2. Waveforms of the emission and the returning echoes. Three targets are located at $(r, \theta, \phi) = (0.730 \text{ m}, 0^\circ, 0^\circ)$, $(1.000 \text{ m}, 0^\circ, 0^\circ)$, and $(1.040 \text{ m}, 0^\circ, 0^\circ)$. Here r is a targets distance from the mouth, θ is its azimuthal angle, and ϕ is its elevational angle. The delay times for the targets are $4295 \mu\text{s}$, $5883 \mu\text{s}$, and $6118 \mu\text{s}$, and the distances from the mouth to each ear via the targets are 1.460 m , 2.000 m , and 2.080 m . The difference between delay times of the second and third targets is $235 \mu\text{s}$, corresponding to 80 mm along the range axis. The EEDNF is found to be 40 kHz for each target. Echoes from the three targets overlap with one another.

be modeled by a bandpass filter bank followed by an envelope-smoothing process according to the SCAT model [13]. In this model, this bandpass filter bank is composed of the 101 constant-bandwidth, second order IIR filters. The center frequencies of these filters are spaced from 20 to 120 kHz and are positioned with regular separations. The envelope-smoothing process is modeled by a half-wave rectification and a low-pass filtering of the output of each the 101 bandpass filters [13, 19–21]. Echoes from the second target and the third target interfere and are not discriminated from each other along the time axis. Figure 3 shows the output of cochlear block for the emission and the echoes of Figure 2.

2.2. THE TEMPORAL BLOCK

The function of the temporal block is to determine the delay times corresponding to the onset and offset in a group of echoes. Thus the distance to the nearest target can be directly determined from the delay time corresponding to onset, and the distance to the farthest target can be directly determined from the delay time corresponding to offset. The processing of the temporal block is made up of the three stages. The first stage is the transform from the envelope pattern into the onset and offset patterns at each bandpass filter channel. The second stage is the integration processing among bandpass filter channels to improve the temporal precision. The third stage is the cross-correlation function between the emission and the echoes returning from targets.

The onset and offset patterns are obtained from the envelope pattern, which is the output from the cochlear block (Figure 4). To obtain the onset and offset patterns, this envelope pattern is transformed into the first derivative pattern by

Figure 3. The spectrogram of the emission and the returning echoes in the case of Figure 2 exhibits an array of smooth envelopes in the case of a linear frequency modulated (LFM) input signal (an FM sweep from 130 kHz to 10 kHz in 3.6 ms).

Figure 4. Determination of the distance to the target region. The extraction of the onset and offset patterns: The envelope pattern in a bandpass filter channel (center frequency = 40 kHz) (upper left), the envelope pattern transformed into the first derivative of the envelope pattern using the sine-wave asymmetric filter (lower left), the onset pattern extracted from the positive part of the first derivative (upper right), and the offset pattern extracted from the negative part of the first derivative are displayed (lower right).

deconvolution of a sine-waveform asymmetric filter with a period of 350 μs . This asymmetric filter is represented by

$$W_{\text{asym}}(t) = \sin(2\pi t/T) \quad 0 < t < T \quad (2)$$

where t is time and T is the period ($= 350 \mu\text{s}$). The onset pattern (B_{on}) and the offset pattern (B_{off}) are the positive part and the negative part of the derivative pattern, respectively. This process has no effect on the determination of the delay time because of deconvoluting both the emission's and echo's envelope patterns in the same way.

In order to obtain the precise onset and offset time, we integrate the onset and offset patterns of different bandpass filter channels by compensating sweep rate. In this model, an 18 kHz integrated filter channel bandwidth is obtained by integrating over 11 bandpass filter channels. The Q_{10dB} values range from 1.4 at 115 kHz to 6.3 at 25 kHz. This integrated filter (IF) in the n th frequency channel is described by the equations

$$I F_{\text{on}}(f_n, t) = \prod_{i=-5}^5 B_{\text{on}}(f_n + FD \times i, t - SR \times i) \quad (3)$$

$$I F_{\text{off}}(f_n, t) = \prod_{i=-5}^5 B_{\text{off}}(f_n + FD \times i, t - SR \times i) \quad (4)$$

where f_n (kHz) is the center frequency for the n th integrated filter channel and t is time (in seconds). The quantity FD ($= 1 \text{ kHz}$) is the frequency difference between bandpass filter channels.

The transform of the time axis into the range axis is carried out by the coincidence circuit, which calculates the cross-correlation function between the emission and the echoes. This cross-correlation function for the n th integrated filter channel is represented by the following equations:

$$C_{\text{on}}(f_n, \tau_i) = \int I F_{\text{on}}(F_n, t + \tau_i) I F_{\text{on}}(f_n, t) dt \quad (5)$$

$$C_{\text{off}}(f_n, \tau_i) = \int I F_{\text{off}}(F_n, t + \tau_i) I F_{\text{off}}(f_n, t) dt \quad (6)$$

$$\tau_i = \tau_0 + i d\tau \quad i = 0, 1, 2, 3, \dots, 70, \quad (7)$$

where the value τ_i is the delay time corresponding to the i th coincidence-detecting cell, and the quantity $d\tau$ is the difference between delay times corresponding to consecutive coincidence-detecting cells ($d\tau = 60 \mu\text{s}$).

The final outputs of the temporal block (O_{on} and O_{off}) are calculated by summing up the outputs of 19 integrated filter channels at each delay time:

$$O_{\text{on}}(\tau_i) = \sum_{f_n} C_{\text{on}}(f_n, \tau_i) \quad (8)$$

$$O_{off}(\tau_i) = \sum_{f_n} C_{off}(f_n, \tau_i) \quad (9)$$

The difference between the onset and offset delay times corresponding to the distances to the nearest and farthest targets, for a group of the echoes returning from single or multiple nearly equidistant targets are determined from O_{on} and O_{off} .

2.3. THE SPECTRAL BLOCK

The function of the spectral block is to determine the differences among delay tunes of echoes from nearly equidistant targets and the EEDNF associated with each target. It indicates that the S_{target} and the S_{ext_ear} should be distinguished from the spectrogram of echoes. It has been known that the EEDNF varies within the somewhat restricted frequency range [15]. While in the EEDNF region the spectrogram includes contributions from both the S_{target} and the S_{ext_ear} , in the non-EEDNF region the spectrogram contains only the S_{target} . Since the echo spectral pattern (S_{echo}) in the non-EEDNF region contains only the interference between echoes from targets, this is transformed into an impulse response to obtain the S_{target} of the whole region and then the S_{ext_ear} can be determined by subtracting the S_{target} from the S_{echo} .

2.4. THE INTEGRATION BLOCK

In the binaural integration block, each target in 3-D space is located using two pieces of information regarding the distance to the target and the EEDNF associated with the target. The distance to the nearest target in the target region can be obtained directly from the delay time corresponding to the onset, which is given as an output of the temporal block. The distance to the farthest target is determined by adding the distance to the nearest target and the distance corresponding to the difference between delay times of the echoes from the two nearly equidistant targets.

3. Results

3.1. MULTIPLE TARGETS WITHOUT INTERFERENCE AMONG ECHOES FROM DIFFERENT TARGETS

As an example of the case of an arbitrary number of targets in 3-D space, we consider that in which there exist five targets and the echoes returning from them overlap but do not interfere with each other. The locations of the five targets are $(r, \theta, \phi) = (0.734 \text{ m}, 6^\circ, 0^\circ), (0.850 \text{ m}, 0^\circ, 0^\circ), (1.064 \text{ m}, 0^\circ, 20^\circ), (1.201 \text{ m}, -8^\circ, 15^\circ),$ and $(1.351 \text{ m}, 2^\circ, 0^\circ)$. The waveforms and spectrograms of the emission and the echoes returning from the five targets are displayed in Figures 5 (a) and (b). We see from these figures that the echoes from the targets overlap but do not interfere with

each other. Figures 5 (c) and (d) display the outputs of the temporal block. The five clear peaks of each curve indicate the delay times for the five targets. These peaks allow us to determine the distances from the mouth to two ears via each target.

In the present case, the echoes from the five targets overlap but do not interfere with each other, so the differences between delay times corresponding to the different targets is on the order of several hundred to several thousand microseconds. Therefore, the echo returning from each target can be easily separated, and the two EEDNFs for each of the five targets can be calculated from the S_{echo} associated with each target. Finally, the direction to each target is determined by using these EEDNFs, as shown in Figure 5 (g). These directions (θ, ϕ) are $(6.0^\circ, 0.0^\circ)$, $(0.0^\circ, 0.0^\circ)$, $(0.0^\circ, 20.0^\circ)$, $(-8.0^\circ, 15.0^\circ)$, and $(2.0^\circ, 0.0^\circ)$. These results are in good agreement with the actual target directions.

3.2. TWO TARGETS WITH THE ALMOST SAME DISTANCE BUT DIFFERENT DIRECTION

We now consider the case in which two targets are located at the almost same distance but different directions. The two targets are located at $(r, \theta, \phi) = (0.734 \text{ m}, 6^\circ, 0^\circ)$ and $(0.768 \text{ m}, -4^\circ, 20^\circ)$. The differences between delay times of the two targets are $189 \mu\text{s}$ (corresponding to 64.3 mm) for the left ear and $209 \mu\text{s}$ (corresponding to 71.1 mm) for the right ear. The EEDNFs corresponding to the nearest target are 43 kHz for the left ear and 37 kHz for the right ear. The EEDNFs corresponding to the farthest target are 34 kHz for the left ear and 38 kHz for the right ear. In this case, we obtain the direction to the nearest target as $(6.8^\circ, 1.0^\circ)$ and to the farthest target as $(-3.2^\circ, 15.0^\circ)$ as shown in Figure 6. There are slight discrepancies between these values and the actual values. However, these discrepancies are acceptable, because the bats appear to determine the direction of an echo's source to within $\pm 2^\circ$ – $\pm 5^\circ$ [17, 22, 23].

3.3. RESOLUTION OF TWO VERY CLOSELY LOCATED TARGETS

We consider the case of two closely located targets at $(r, \theta, \phi) = (0.7300 \text{ m}, 0^\circ, 0^\circ)$ and $(0.7308 \text{ m}, 0^\circ, 0^\circ)$. The distances from the mouth to the ear via the nearest and the farthest target are 1.460 m ($= 4294.1 \mu\text{s}$) and 1.4616 m ($= 4298.8 \mu\text{s}$). Thus, there is the difference between delay time of $4.7 \mu\text{s}$ ($= 1.6 \text{ mm}$). Each of the four EEDNFs corresponding to these targets is 40 kHz. The calculated difference between delay times for the two targets is $4.1 \mu\text{s}$ ($= 1.4 \text{ mm}$). Although this value differs slightly from the true value, our model has successfully determined that there are two nearly equidistant targets. Each of the four EEDNFs is found to be 40kHz, as shown in Figure 7 (f). This implies that the directions of both targets are $(0.0^\circ, 0.0^\circ)$ in agreement with the actual values.

Figure 5. Multiple targets with no interference among echoes from targets. Five targets are located at $(r, \theta, \phi) = (0.734 \text{ m}, 6^\circ, 0^\circ)$, $(0.850 \text{ m}, 0^\circ, 0^\circ)$, $(1.064 \text{ m}, 0^\circ, 20^\circ)$, $(1.201 \text{ m}, -8^\circ, 15^\circ)$, and $(1.351 \text{ m}, 2^\circ, 0^\circ)$. (a) Waveforms of the emission and the returning echoes. (b) Spectrogram of the emission and the echoes. The echoes from the five targets overlap but do not interfere with each other. (c) & (d) Outputs of the temporal block for left and right ears, respectively. The solid and dotted curves represent outputs obtained from the onset and offset patterns, respectively. The five clear peaks of each curve indicate the delay times to the five targets. The distances from the mouth to the two ears for the targets can be determined from this information. (e) The directions of the five targets. The calculated direction (θ, ϕ) is $(6.0^\circ, 0.0^\circ)$ for the first target (solid square), $(0.0^\circ, 0.0^\circ)$ for the second target (open square), $(0.0^\circ, 20.0^\circ)$ for the third target (open triangle), $(-8.0^\circ, 15.0^\circ)$ for the fourth target (solid triangle), $(2.0^\circ, 0.0^\circ)$ for the fifth target (open circle). Each of these calculated directions is in good agreement with the actual direction.

Figure 6. Outputs in the case of two targets at almost equal distances but different directions. The two targets are located at $(r, \theta, \phi) = (0.734 \text{ m}, 6^\circ, 0^\circ)$ and $(0.768 \text{ m}, -4^\circ, 20^\circ)$. The delay times associated with the first target are $4325 \mu\text{s}$ ($= 1.470 \text{ m}$) for the left ear and $4312 \mu\text{s}$ ($= 1.466 \text{ m}$) for the right ear. The delay time associated with the second target are $4514 \mu\text{s}$ ($= 1.535 \text{ m}$) for the left ear and $4521 \mu\text{s}$ ($= 1.537 \text{ m}$) for the right ear. (a) The waveforms of the emission and the returning echoes. (b) Spectrogram of the emission and the echoes. (c) and (d) Outputs of the temporal block for the left ear and the right ear, respectively. The solid curve represents the output calculated from the onset pattern and the dotted curve represents the output calculated from the offset pattern. (c) The delay times corresponding to the peaks of the onset and offset patterns are $4358 \mu\text{s}$ ($= 1.482 \text{ m}$) and $4478 \mu\text{s}$ ($= 1.522 \text{ m}$). (d) The delay times corresponding to the peaks of the onset and offset patterns are $4298 \mu\text{s}$ ($= 1.461 \text{ m}$) and $4538 \mu\text{s}$ ($= 1.543 \text{ m}$). (e) Calculated directions of the two targets. The direction of the nearer target (solid square) is $(6.8^\circ, 1.0^\circ)$, and that of the farther target (gray square) is $(-3.2^\circ, 15.0^\circ)$.

Figure 7. Resolution of two closely located targets. Two targets are located at $(r, \theta, \phi) = (0.7300 \text{ m}, 0^\circ, 0^\circ)$ and $(0.7308 \text{ m}, 0^\circ, 0^\circ)$. The delay times are $4294.1 \mu\text{s}$ ($= 1.4600 \text{ m}$) for the first target and $4298.8 \mu\text{s}$ ($= 1.4616 \text{ m}$) for the second target. The difference between delay times of echoes from the two targets is $4.7 \mu\text{s}$ ($= 1.6 \text{ mm}$). (a) The waveforms of the emission and the echoes. (b) Spectrogram of the emission and the echoes. (c) Output of the temporal block. The delay times corresponding to the peaks of the onset and offset patterns are both $4298 \mu\text{s}$ ($= 1.461 \text{ m}$). (d) $S_{\text{ext_ear}}$ associated with the nearer target for the left ear. The EEDNF is 40 kHz.

Figure 8. Robustness in a noisy environment. The signal-to-noise ratio is 5 dB. Two targets are located at $(r, \theta, \phi) = (0.730 \text{ m}, 0^\circ, 0^\circ)$ and $(0.738 \text{ m}, 0^\circ, 0^\circ)$. The delay time are $4295 \mu\text{s}$ ($= 1.460 \text{ m}$) for the first target and $4342 \mu\text{s}$ ($= 1.476 \text{ m}$) for the second target. The difference between delay times of echoes from the targets is $47 \mu\text{s}$. The EEDNF is 40 kHz for both ears and both targets. (a) The waveforms of the emission and the echoes. (b) Spectrogram of the emission and the echoes. (c) Outputs of the temporal block. The solid curve represents the output calculated from the onset pattern and the dotted curve represents the output calculated from the offset pattern. The delay times corresponding to the peaks of the onset and offset patterns are both $4298 \mu\text{s}$ ($= 1.461 \text{ m}$). (d) $S_{\text{ext_ear}}$ associated with the nearer target. The EEDNF for the nearer target is 40 kHz.

3.4. ROBUSTNESS IN A NOISY ENVIRONMENT

In this study, two targets are located at $(r, \theta, \phi) = (0.730 \text{ m}, 0^\circ, 0^\circ)$ and $(0.738 \text{ m}, 0^\circ, 0^\circ)$. The rms amplitude signal-to-noise ratio for a single target is 5.0 dB. The distances from the mouth to the ear via the nearest and farthest targets are 1.460 m ($= 4295 \mu\text{s}$) and 1.476 m ($= 4342 \mu\text{s}$). The difference between delay times for the two targets is $47 \mu\text{s}$ ($= 16 \text{ mm}$). All EEDNFs for the two targets and two ears are 40 kHz. From the outputs of the temporal block, the delay time for the onset and the offset of the group of echoes are found to be $4298 \mu\text{s}$ ($= 1.461 \text{ m}$) and $4298 \mu\text{s}$ ($= 1.461 \text{ m}$) for both ears, as shown in Figure 8(c). Figure 8 (d) displays the selective-pass impulse response ($I_{\text{sel_pass}}$). The calculated difference between delay times is $46.8 \mu\text{s}$ ($= 15.9 \text{ mm}$), which is in good agreement with the actual difference between delay times. Figure 8 (e) displays the $S_{\text{ext_ear}}$ associated with the nearest target for the left ear. The EEDNF calculated is 40 kHz. The EEDNF associated with each target for each ear was similarly calculated and found to be 40kHz, in good agreement with the actual values. These results indicate that the direction to each target is $(0.0^\circ, 0.0^\circ)$. We thus find that in spite of the existence of noise, the locations of the targets can be accurately determined.

4. Discussion and Conclusion

When multiple targets are located at distances which differ enough that the differences between the delay times of their returning echoes are longer than the duration of the emission, 3.6 ms, the correspondence between echoes and targets can be easily discriminated and these echoes can be processed independently to locate each target. The difficult problem is to locate multiple targets whose distances are close enough that differences between the delay times of their echoes are shorter than the duration of the emission. In this case, the echoes from the multiple targets overlap. This situation can be classified into two cases, that in which the difference between delay times is longer than the integration time of the bandpass filter, and that in which it is not. In the cochlear block, the waveform of the returning echoes are transformed into the spectrogram by the bandpass filter, the half-rectification, and the lowpass filter with an integration time of about $350 \mu\text{s}$. When the difference between delay times of echoes is longer than the integration time of the filter, each envelope pattern in the bandpass filter channel can be separated along the time axis. In this case, the distance and the EEDNF for each target can be calculated from each separated S_{echo} . As an example of such a situation, five targets were successfully located, as shown in Figure 5.

By contrast, when the targets are so close that the difference between their corresponding delay times is shorter than the integration time, within each filter channel the envelope patterns produced by these nearly equidistant targets cannot be separated along the time axis. The difference between delay times of echoes from nearly equidistant targets is determined from the S_{echo} associated with the average of the delay time corresponding to the onset and the offset of the group of

interfering echoes. The S_{echo} associated with the average delay time is calculated from the cross-correlation function ($C(\tau_i, f_n)$). We divide the S_{echo} into two regions, the EEDNF and the non-EEDNF regions, along the frequency axis. The S_{echo} in the EEDNF region consists of both the S_{target} and the $S_{\text{ext_ear}}$ while the S_{echo} in the non-EEDNF region contains only S_{target} . The difference between delay times of the echoes returning from nearly equidistant targets can be accurately determined from the $I_{\text{sel_pass}}$. Since the $I_{\text{sel_pass}}$ is not influenced by the interference caused by the physical process of reception at the external ear, the S_{target} in the entire frequency region can be constructed by using the $I_{\text{sel_pass}}$ of the S_{echo} associated with the average delay time in the non-EEDNF region. In a noisy environment, we found that the difference between delay times can be accurately determined with a signal-to-noise ratio of 5 dB (Figure 8 (d)). Since the difference between delay times can be accurately determined with an error of less than $1 \mu\text{s}$ through the deconvolution of the selective-pass improved basis vectors, the calculated difference between delay times can be used as crucial information for the determination of the azimuthal direction.

In the temporal block, the envelope pattern can be transformed into the onset and offset patterns using a sine-wave asymmetric filter. The onset and offset patterns both have a peak, the former at the time corresponding to the beginning of the envelope pattern and the latter at the time corresponding to the end, because the transform extracts the derivative of the envelope pattern, with ripples removed. Also, the integrated patterns have a sharp peak along the time axis. While the width of both the onset and offset patterns at 50% of peak height are approximately $150 \mu\text{s}$ along the time axis, for the integrated patterns, these widths are approximately $30 \mu\text{s}$. Thus, the integrated onset and offset patterns allows for more accurate determination of the delay times, because the width of the peaks is reduced by a factor of about one-fifth. This width at half peak height along the time axis equals the standard deviation of the spike latency of phasic neurons in the VNLLc [24]. Therefore, it seems possible that FM bats carry out something equivalent to integration processing in bandpass filter channels. Also, we used a value of $60 \mu\text{s}$ for the time interval between coincidence-detecting cells, because this value is close to that found for the interval between BDs of delay-tuned neurons in bats, $100 \mu\text{s}$ [25]. Thus, in our model, the distances to targets can be determined within a precision of $60 \mu\text{s}$ by using only the delay times corresponding to the peaks of the onset and offset patterns. To determine the distances to the targets producing interfering echoes at higher resolution requires interpolation or population coding.

In this model, the most important mechanism is the method to discriminate the S_{target} and the $S_{\text{ext_ear}}$ from the S_{echo} . Since bats can capture the moving target in the presence of multiple objects, it is thought that this method is actually used by FM bats in the behavioral studies. However, it has been unknown that the S_{target} and the $S_{\text{ext_ear}}$ are represented in the bat's brain in the physiological study. For this reason, it requires to verify the existence of the neurons, which can respond only the S_{target} or the $S_{\text{ext_ear}}$. The existence of neurons, which can respond only the

S_{target} , may be verified by following two experiments. The first one is to indicate that the responses of neurons depend on the difference between delay times when the difference between delay times is altered and directions of these targets are not altered. The second one is to indicate that the responses of neurons do not change when the difference between delay times is not altered and directions of these targets are altered. Contrastingly, the existence of neurons which can respond only the $S_{\text{ext_ear}}$ may be verified by two experiments in the same way. Also, it requires the external ear's transfer function to verify the existence of these neurons.

Developing this study, it will be possible to detect the positions and shapes of plural objects in a dark and complex environment such as the deep sea and the underground space. The sonar image made by a single emission has a great advantage to recognize such environments. Especially, the land mines laid in the world are the serious problem. However, it is promising to solve the land mine problem by establishing the echolocation system, as the bats perceive the external world.

Acknowledgement

This study is supported by 'Research for the Future' Program (JSPS_RFTF97I00103) of Japan Society for the Promotion of Science.)

References

1. Griffin, D.R.: *Listening in the dark*, Yale U.P., New Haven CT, 1958, reprinted by Cornell U.P., Ithaca, New York, 1986.
2. Simmons, J.A., Ferragamo, M.J., Saillant, P.A., Haresign, T., Wotton, J.M., Dear, S.P. and Lee, D.N.: Auditory Dimensions of Acoustic Images in Echolocation, In: A.N. Popper and R.R. Fay (eds.), *Hearing by Bats*, Springer, New York, 1995, pp. 146–190.
3. Dear, S.P., Simmons, J.A. and Fritz, J.: A Possible Neuronal Basis for Representation of Acoustic Scenes in Auditory Cortex of the Big Brown Bat, *Nature* **364** (1993), 620–623.
4. Dear, S.P., Fritz, J., Haresign, T., Ferragamo, M. and Simmons, J.A.: Tontopic and Functional Organization in the Auditory Cortex of the Big Brown Bat *Eptesicus fuscus*, *J. Neurophysiol.* **70** (1993), 1988–2009.
5. Kawasaki, M., Margoliash, D. and Suga, N.: Delay-Tuned Combination-Sensitive Neurons in the Auditory Cortex of the Vocalizing Mustached Bat, *J. Neurophysiol.* **59** (1988), 623–635.
6. Suga, N., O'Neill, W.E., Manabe, T.: Cortical Neurons Sensitive to Particular Combinations of Information-Hearing Elements of Biosonar Signals in the Mustached Bat, *Science* **200** (1978), 778–781.
7. Sullivan, W.E.: Neural Representation of Target Distance in Auditory Cortex of the Echolocating Bat, *Myotis lucifugus*, *J. Neurophysiol.* **48** (1982), 1011–1032.
8. Sanderson, M.I. and Simmons, J.A.: Neural Responses to Overlapping FM Sounds in the Inferior Colliculus of Echolocating Bats, *J. Neurophysiol.* **83** (2000), 1840–1855.
9. Simmons, J.A.: The Resolution of Target Range by Echolocating Bats, *J. Acoust. Soc. Am.* **54** (1973), 157–173.
10. Miller, L.A.: Arctiid Moth Clicks can Degrade the Accuracy of Range Difference Discrimination in Echolocating Big Brown Bats, *Eptesicus fuscus*, *J. Comp. Physiol. A* **168** (1991), 571–579.

11. Surlykke, A.: Target Ranging and the Role of Time-Frequency Structure of Synthetic Echoes in Big Brown Bats, *Eptesicus fuscus*, *J. Comp. Physiol. A* **170** (1992), 83–92.
12. Simmons, J.A., Ferragamo, M., Moss, C.F., Stevenson, S.B. and Altes, R.A.: Discrimination of Jittered Sonar Echoes by the Echolocating Bat *Eptesicus fuscus*: The Shape of Target Images in Echolocation, *J. Comp. Physiol. A* **167** (1990), 589–616.
13. Saillant, P.A., Simmons, J.A., Dear, S.P. and McMullen, T.A.: A Computational Model of Echo Processing and Acoustic Imaging in Frequency-Modulated Echolocating Bats: The Spectrogram Correlation and Transformation Receiver, *J. Acoust. Soc. Am.* **94**, (1993) 2691–2712.
14. Fuzessery, Z.M.: Monaural and Binaural Spectral Cues created by the External Ears of the Pallid Bat, *Hea. Res.* **95** (1996), 1–17.
15. Wotton, J.M., Haresign, T. and Simmons, J.A.: Spatially Dependent Acoustic Cues generated by the External Ear of the Big Brown Bat, *Eptesicus fuscus*, *J. Acoust. Soc. Am.* **98** (1995), 1423–1445.
16. Wotton, J.M., Haresign, T., Ferragamo, M.J. and Simmons, J.A.: Sound Source Elevation and External Ear Cues Influence the Discrimination of Spectral Notches by the Big Brown Bat, *Eptesicus fuscus*, *J. Acoust. Soc. Am.* **100** (1996), 1764–1776.
17. Wotton, J.M. and Simmons, J.A.: Spectral Cues and Perception of the Vertical Position of Targets by the Big Brown Bat, *Eptesicus fuscus*, *J. Acoust. Soc. Am.* **107** (2000), 1034–1041.
18. Simmons, J.A., Saillant, P.A., Wotton, J.M., Haresign, T., Ferragamo, M.J. and Moss, C.F.: Composition of Biosonar Images for Target Recognition by Echolocating Bats, *Neural Networks* **8** (1995), 1239–1261.
19. Brugge, J.F., Andersom D.J., Hind, J.E. and Rose, J.E.: Time Structure of Discharges in Single Auditory Verve Fibers of the Squirrel Monkey in Response to Complex Periodic Sounds, *J. Neurophysiol.* **32** (1969), 386–401.
20. Rose, J.E., Brugge, J.F., Andersom D.J. and Hind, J.E.: Phase-Locked Responses to Low-Frequency Tones in Single Auditory Verve Fibers of the Squirrel Monkey, *J. Neurophysiol.* **30** (1967), 769–793.
21. Russen, I. and Palmer, A.: Filtering due to the Inner Hair-Cell Membrane Properties and its Relation to Phase-Locking Limit in Cochlear Verve Fibres, In: B.C.J. Moore and R.D. Patterson (eds.), *Auditory Frequency Selectivity*, Plenum, New York, 1986, pp. 199–206.
22. Grinnell, A.D.: Hearing in Bats: An Overview, In: A.N. Popper and R.R. Fay (eds.), *Hearing by Bats*, Springer, New York, 1995, pp. 1–36.
23. Masters, W.M., Moffat, A.J.M. and Simmons, J.A.: Sonar Tracking of Horizontally Moving Targets by the Big Brown Bat *Eptesicus fuscus*, *Science* **228** (1985), 1311–1333.
24. Covey, E. and Casseday, J.H.: The monaural nuclei of the lateral lemniscus in an echolocating bat: parallel pathways for analyzing temporal features of sound, *J. Neurosci.* **11** (1991), 3456–3470.
25. Suga, N. and O'Neill, W.E.: Neural axis representing target range in the auditory cortex of the mustached bat, *Science* **206** (1979), 778–791.