euronoise

A model of range discrimination of multiple objects by using the linear period modulation signal

Ikuo Matsuo

Tohoku Gakuin University, Tenjinzawa 2-1-1, Izumi-ku, 9813193 Sendai, Japan
matsuo@cs.tohoku-gakuin.ac.jp

Using the echolocation, bats can capture moving objects in 3D space. The big brown bats, *Eptesicus fuscus*, emit the linear period modulation sound and can identify objects with an accuracy of less than a millimeter. The delay separation of three or more closely spaced objects can be determined through analysis of the echo spectrum. However, delay times of these objects cannot be properly estimated by using only the echo spectrum because the temporal sequence of delay separations cannot be determined without information on temporal changes in the interference pattern of the echoes. We previously proposed the model to determine delay times of multiple objects from the echoes by using the linear frequency modulation sound. In order to extract the temporal changes, Gaussian chirplets with a carrier frequency compatible with the emission sweep rates were introduced. In the case of the linear period modulation sound, the method to determine delay times of multiple objects from echoes, however, has been never proposed. In this paper, the model to determine delays of multiple objects is proposed by introducing the chirplet filters, which can extract the temporal changes dependent on delay times of multiple objects.

1 Introduction

Bats emit trains of high-frequency sound waves and can locate an individual object among three or more objects using the temporally overlapping echoes of the emitted waves [1, 2]. Experimental evidence indicates that FM bats are capable of locating objects with sub microsecond accuracy [3, 4]. For two or more closely spaced objects, the echoes will interfere with one another. Convolution of the constant frequency (CF) carrier wave with the echoes allows delay separations among them to be estimated from the spectrum or spectrogram of echoes [5, 6]. It is, however, difficult to determine the temporal sequence of delay separations since this transformation cannot extract temporal changes in the echo spectrum.

To accurately locate three or more objects it is necessary to extract such temporal changes in the echo spectrum. First, it is necessary to accurately determine the delay time T_1 for the first object, because the error in T_1 seriously affects the determination of the temporal sequence of delay separations. It has been found that bats can measure each delay time T_i for two closely spaced objects with μs accuracy [2, 7]. T_1 can be approximated using the cross-correlation function of the spectra of emitted and reflected waves [5]. However, it is difficult to accurately determine T_1 from this cross-correlation function because the integration time of cochlear filters is large (several hundreds of microseconds). For the accurate determination of T_1, a new method was proposed which utilizes the echo spectrum [8]. The second requirement for the determination of the temporal sequence of delay separations is the extraction of temporal changes in the interference patterns from the echo spectra. It is difficult to extract these temporal changes through the application of the usual method of convolving the echoes with the CF carrier wave to transform echoes into the echo spectra, because this method is relatively insensitive to temporal changes. In the case of the linear frequency modulation sound, the Gaussian chirplets for which the carrier frequency agrees with the sweep rate of this emission were introduced, and the method for the accurate determination of each delay time T_i for three or more closely spaced objects was proposed [8]. In the case of the linear period modulation sound that bats, *Eptesicus fuscus*, emit, the method to determine delays of multiple objects from echoes, however, has been never proposed. In this paper, the model to determine delays of multiple objects is proposed by introducing the chirplet filters, which can extract the temporal changes dependent on delay times of multiple objects.

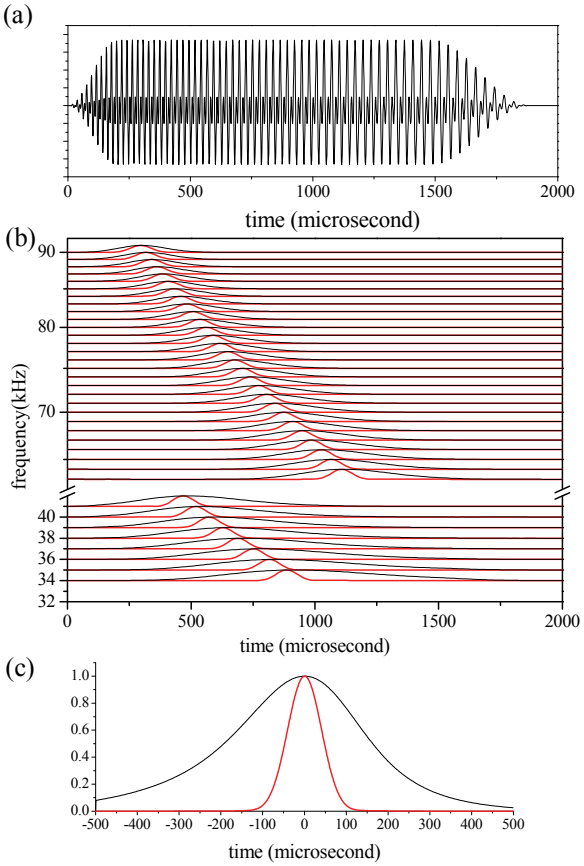

Fig.1. (a) Emission waveform. (b) Emission spectrogram. (c)Temporal emission pattern W_{emi} corresponding to the spectrogram in one bandpass filter.

2 Model

2.1 Characteristics of emission and echo

Bats, *Eptesicus fuscus*, emit linear period modulation during echolocation. This emission consists of several harmonics and ranges from about 20 kHz to one hundred kHz [9]. This model uses an emission with first and second harmonics. The frequency of the first harmonics begins at 53 kHz and sweeps down to 25 kHz. The duration of the emission is about 1.9 ms, and the rise and fall time are 200 microsecond. Figure 1 (a) shows the emitted waveform.

In this paper, the following conditions are assumed to simplify the simulation of the echo. Each object is a point object. An extended object is defined as consisting of a distribution of acoustically reflecting points, called “glints.” Reflection does not modify the phase of the sound wave.

The plot of reflected intensity vs. distance (equivalent to delay time) is called the "reflected intensity distribution."

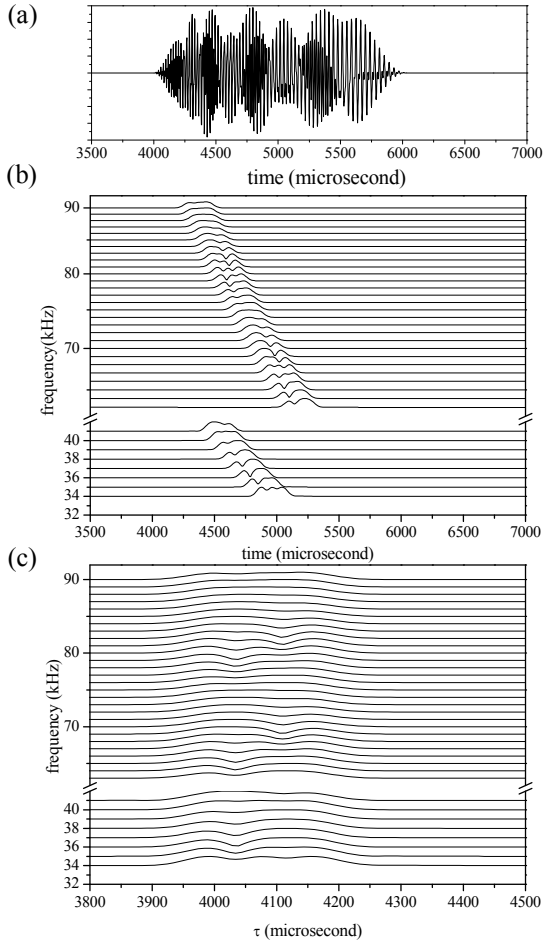

Fig.2. (a) Echo waveform. (b) Echo spectrogram. (c) Range -frequency pattern computed from the spectrogram.

2.2 Transformation of the waveform into the spectrogram by chirplet filters

Waveforms for emission and echoes from many closely spaced objects were put into this model. These waveforms were transformed into spectrograms in a manner that simulates this process in the mammalian cochlea. The carrier wave previously used for cochlear bandpass filters was a wave of a constant frequency (CF) corresponding to the center frequency of the bandpass filter [5, 6], which is called "tone" filter. It was previously impossible to extract temporal changes in the interference patterns of echoes due to the convolution of the CF carrier wave with the both the FM wave emitted and the echoes. In order to extract temporal changes in the interference pattern, a chirplet filter, which has Gaussian amplitude characteristic and phase characteristic which is dependent on the sweep rate of emission, is introduced. This bandpass filter bank is composed of 76 filters. The center frequencies of these filters range from 25 to 100 kHz, positioned at regular frequency intervals. The 10 dB frequency bandwidth of each filter is 12 kHz. Q_{10dB} values range from 2.1 at 25 kHz to 8.3 at 100 kHz. The waveforms for both the emitted waves and echoes were transformed into the spectrogram P (f, t) by convolution with this filter. The red line in Fig. 1 (b) and (c) shows the outputs of the chirplet filters and black line shows the outputs of the tone filters for the emitted waveform. Figure 1 (c) shows the temporal pattern corresponding to the spectrogram of emission in one filter (with center frequency 40 kHz). This temporal pattern is termed "W_{emi}." This model uses only filter's outputs at the center frequency (34 – 41 kHz and 63 – 90 kHz) whose shapes of temporal patterns W_{emi} are same with one another. As shown in Fig. 1 (c), the temporal changes in the interference patterns can be extracted by convolution of the chirplets.

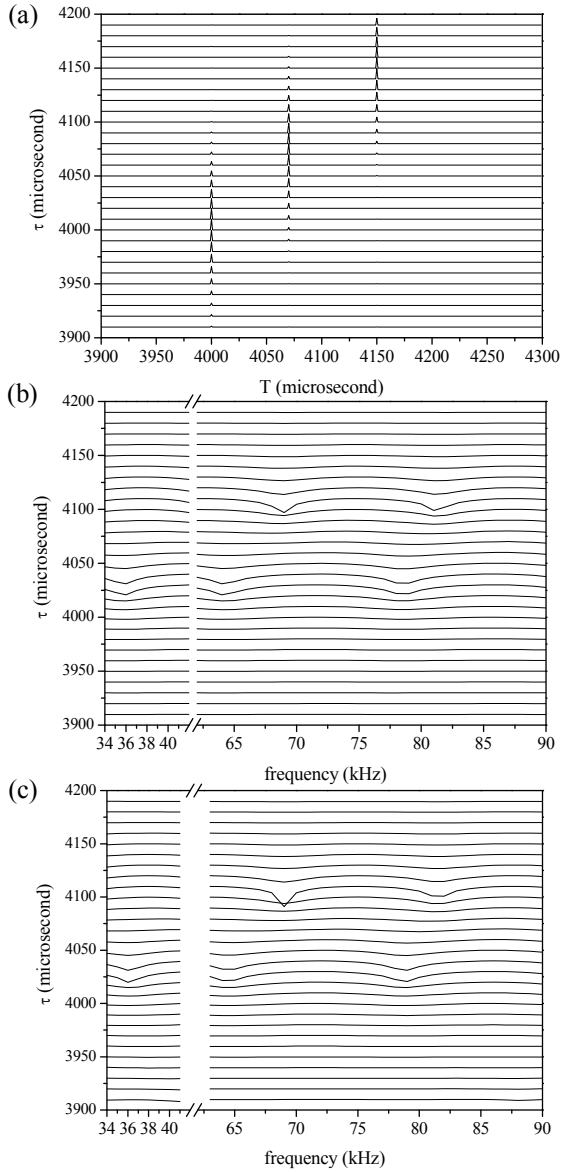

Fig.3. Correspondence of the output from the cochlear filters with the reflected intensity distribution. (a) Windowed reflected intensity distribution R. (b) Range frequency pattern S_{ref} computed from the reflected intensity distribution. (c) Range frequency pattern S_{echo} computed by convolution with chirplets.

In order to demonstrate the output from cochlear filters for echoes originating from many closely spaced objects, we considered the situation of three such objects. The delay times for these objects were 4000, 4070 and 4150 μs and the reflectivity of each object was 1.00. Figures 2 (a) and (b) display the waveforms of the echoes and the output of the chirplets, respectively. This spectrogram P (f, t) was transformed into a range-frequency pattern S_{echo} (f, τ) with 10 μs interval by compensating for the sweep rate of the

emission frequency modulation. Since the start time for emission was 0, this compensating time was represented by τ, which implies the range corresponding to the delay time. Figure 2 (c) represents the range-frequency pattern computed using chirplet filters.

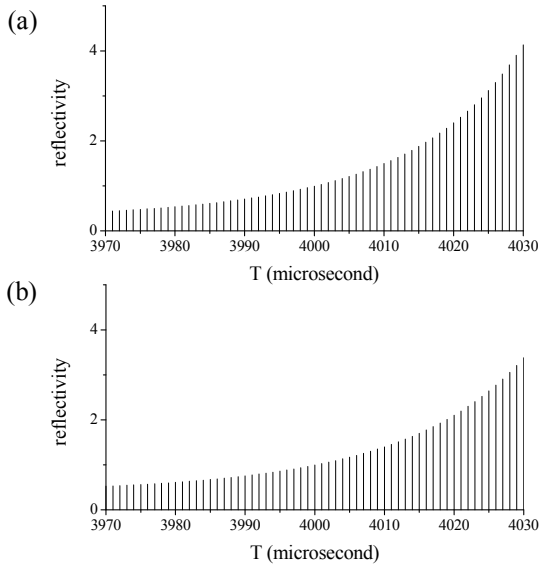

Fig.4. Determination of T_1.

2.3 Correspondence of the output of cochlear filters with the intensity distribution of reflections along the range axis

In order to demonstrate the correspondence of the outputs of cochlear filters with the reflected intensity distribution along the range axis (T), range-frequency patterns S_{ref} (f, τ) were theoretically computed from the reflected intensity distribution. This range-frequency pattern, computed by multiplying the reflected intensity distribution by values of W_{emi} whose peak time corresponds to τ, is called the "windowed reflected intensity distribution," R (T, τ). Figure 3 (a) shows the windowed reflected intensity distribution R that is represented along the range axis at each delay τ. S_{ref} (f, τ) was computed by Fourier transforming the windowed reflected intensity distribution in Fig. 3 (b). As shown in Figs. 3 (b) and (c), the spectral shape of the S_{echo} (f, τ) closely corresponds to that of the S_{ref} (f, τ) at each τ.

2.4 The determination of delay times for closely spaced objects

Determination of each delay time for three or more closely spaced objects requires estimation of T_1 and the temporal sequence of delay separations from the range-frequency pattern. T_1 was determined from the echo spectra around the onset delay τ_{on} since here, these are mainly influenced by the echo of the first object. T_1 and the corresponding reflectivity r_1 are uniquely determined from the averages of the two spectra at the onset delay time τ_{on} and 10 μs later, because the average of S_{ref} at an arbitrary τ is dependent upon both T_1 and r_1. As shown in Fig. 4 (a), candidates for T_1 from the reflected intensity distribution are estimated from the average of S_{echo} at the onset delay τ_{on} (3940 μs). Figure 4 (b) shows candidates estimated from the average of S_{echo} 10 μs after τ_{on} (3950 μs). Finally, the delay time T_1 is determined as 4000 μs by estimation of the correspondence between reflected intensities of the two averages of S_{echo} at τ = 3940 μs and τ = 3950 μs.

T_2 was determined by adding T_1 to the estimated delay separation between the first and second objects. Further delay times T_i were determined by iterating these procedures. In the case of many closely spaced objects, S_{echo} at the current delay τ includes interference patterns dependent on the delay separations between these objects. Therefore, it was necessary to remove interference patterns dependent on the delay separations estimated at the previous delay τ. In order to exclude these interferences, S_{ref} at the current delay was computed from the reflected intensity distribution determined at the previous delay. The interference pattern dependent upon the next delay separation was computed by subtracting S_{ref} from S_{echo}, as S_{ref} from closely spaced objects corresponds to the summation of interference patterns dependent on delay separations between these objects. This subtracted spectrum is labeled S_{sub}.

The cepstrum was computed from S_{sub} by inverse Fourier transform in order to estimate the delay separations for these objects. Since the frequency range was 34 to 41 kHz and 63 to 90 kHz in this model, the spectrum, which has constant amplitude along the frequency axis, was transformed into the cepstrum, with sidelobes except at t=0 through deconvolution of cosine and sine functions. In order to eliminate the sidelobes, these cosine and sine functions were modified into W_{cos} and W_{sin}, respectively [5, 8].

$$W_{\cos}(f_n,t) = \cos(2\pi t f_n) - \left\{ \sum_f \frac{\cos(2\pi t f)}{\mathrm{N}} \right\}$$

$$W_{\sin}(f_n,t) = \sin(2\pi t f_n) - \left\{ \sum_f \frac{\sin(2\pi t f)}{\mathrm{N}} \right\}$$

where f_n is the center frequency of the cochlear filter and N is the number of cochlear filters. Using the deconvolution of W_{cos} and W_{sin}, S_{sub} was transformed into the cepstra, C_{cos} (t, τ) and C_{sin} (t, τ), respectively. The C (t, τ) were computed using the equation:

$$C(t,\tau) = \sqrt{(C_{\cos}(t,\tau))^2 + (C_{\sin}(t,\tau))^2}.$$

Determination of the next delay separation from S_{sub} requires estimation of the intensity of the periodic component in the cepstrum C. Since this intensity is represented by $2r_1r_2$ ($r_1 \geq r_2$), $2r_1r_2$ was computed by extracting the peak of the cepstrum C_{peak}. This value is dependent on both r_1 and r_2. Estimation of the appearance of interference independent of each reflectivity requires the estimation of the correspondence between the reflectivity of two objects. Here, we used the value of C_{peak} divided by C at t=0 as the criterion for estimation of appearance of the interference. In the case of two objects, this value corresponds to $2r_1r_2/(r_1^2+r_2^2)$. When this value was more than 0.5, corresponding to almost 4.8 dB of the difference between the peak and trough amplitudes in the logarithm spectrum, the delay separation was estimated from S_{sub}.

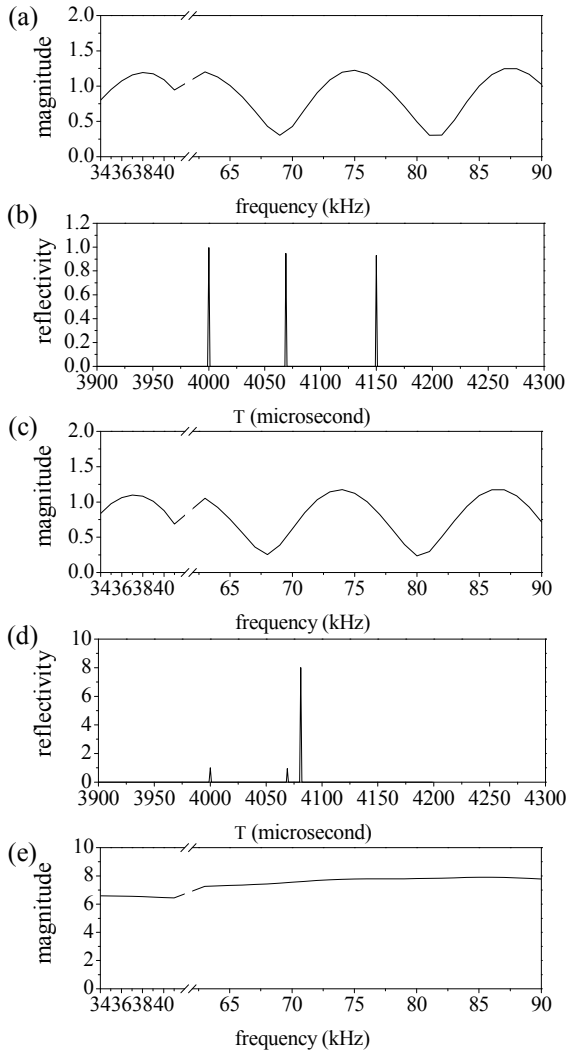

Fig.5. Determination of the temporal sequence of the delay separations. To find T_3, two candidates in the reflected intensity distribution are estimated from the delay separation determined from S_{echo} at delay τ. (a) S_{echo} at delay time τ. (b) Estimated delay separation corresponds to that between the second and third objects. (c) S_{ref} computed from the candidate from the reflected intensity distribution in (b). (d) Estimated delay separation corresponds to that between the first and third objects. (e) S_{ref} computed from the candidate in the reflected intensity distribution in (d).

The determination of each T_i for three or more closely spaced objects requires the determination of the temporal sequence of delay separations. In the process of estimating each delay time T_i, there arise several candidates from the reflected intensity distributions. In the determination of T_3, the estimated delay separation either corresponds to the delay separation between the first and third objects or that between the second and third objects. Here it was therefore necessary to estimate two candidates for T_3. For each, the reflectivity r_3 was computed using the value of C_{peak}, which was in agreement with the values calculated by multiplying values from the windowed reflected intensity distribution at the delay times found for these two objects. Figures 5 (b) and (d) show two candidates from the reflected intensity distributions. In order to determine T_3, S_{ref} at the current delay τ was computed from each candidate of the reflected intensity distribution (Fig. 5 (c), (e)). When the estimated reflected intensity distribution corresponds to the actual reflected intensity, S_{ref} corresponds to S_{echo} at each delay τ in Fig. 5 (a). Therefore, the reflected intensity distribution was determined by comparing S_{echo} with S_{ref} at this delay τ. T_i for the second and third objects were determined to be 4069 and 4150 μs.

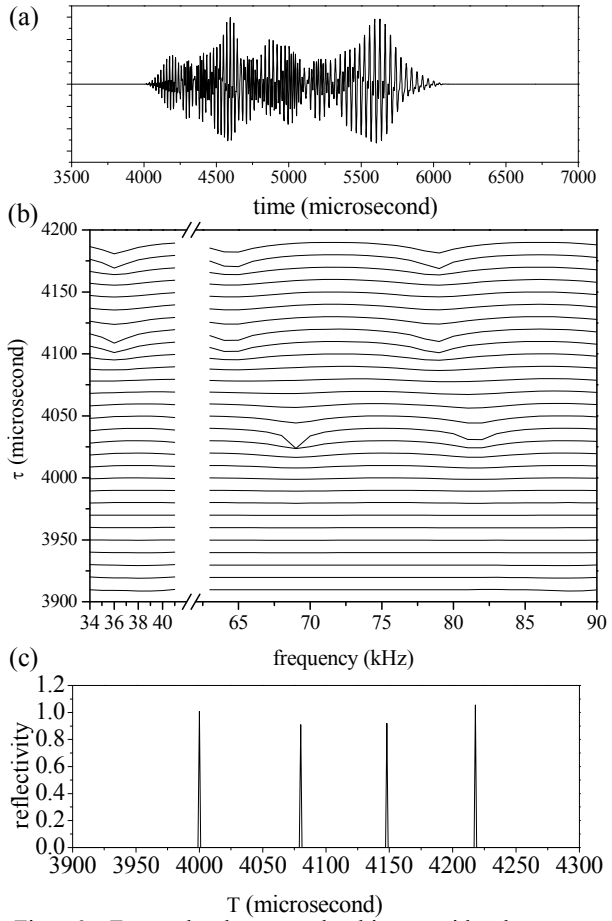

Fig. 6. Four closely spaced objects with the same reflectivity.

3 Results

3.1 Four closely spaced objects with the same reflectivity

As an illustration of the case for an arbitrary number of objects, we considered a situation with four objects having the same reflectivity. T_i was 4000, 4080, 4150 and 4220 μs, and the reflectivity for these objects was 1.00. The waveforms and range-frequency pattern of the echoes returning from these objects are displayed in Figs. 6 (a) and (b). T_i for the four objects were determined to be 4000, 4080, 4148 and 4218 μs and their reflectivities were determined to be 1.01, 0.91, 0.92, and 1.06 respectively, from the reflected intensity distribution in Fig. 6 (c). This is in very good agreement with the actual delay times of the objects while the reflectivities are in reasonable agreement.

3.2 Three closely spaced objects with the different reflectivity

We considered the case of three objects with different reflectivities. T_i for the objects was 4000, 4050 and 4130 μs and their reflectivities were 1.00, 0.20 and 1.00, respectively. The waveforms of the echoes from the four objects are displayed in Figs. 7 (a). T_i for the objects were

determined as 4000, 4050 and 4130 μs and the reflectivities were found to be 1.01, 0.20 and 0.97 respectively, from the reflected intensity distribution in Fig. 7 (b). The calculated delay times are in very good agreement with the actual delay times and reflectivities were determined fairly accurately with this model.

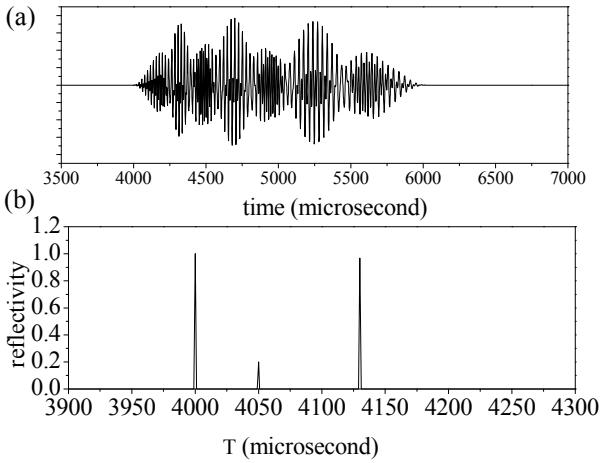

Fig. 7. Three closely spaced objects with the different reflectivity.

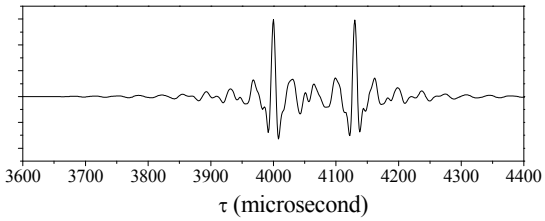

Fig.8. Cross-correlation function between the emitted and reflected waveform for three objects with the different reflectivity..

4 Conclusion

In the case of many closely spaced objects, the echoes interfere with one another. In order to determine each T_i it is necessary to solve two problems: how to accurately determine T_1, and how to determine the temporal sequence of the delay separations. The delay time T_1 of the first object can be determined from the echo spectra around the onset. In this model, T_1 was determined from the echo spectrum at the onset delay τ_{on} and the spectrum 10 μs later. This indicates that the T_1 can be estimated independent of the integration time of the cochlear filters and can be determined from two echo spectra with an accuracy of a microsecond. The delay time T_2 of the second object can be determined by adding T_1 and the estimated delay separation between object 1 and 2. Further objects can be located in sequence by this procedure. As shown in Fig. 6 and 7, this model can determine delay times for three or more closely spaced objects with an accuracy of about several microseconds.

It should be noted that there is a large discrepancy between the delay times calculated using this method and those calculated using the cross-correlation function between the waveforms of emission and reflection. The delay times found using the latter method are determined from the peak times of the cross-correlation function. Figure 8 shows the cross-correlation function in the case of three objects with different reflectivity. The delay times for these objects are 4000, 4050 and 4130 μs and reflectivities are 1.00, 0.20 and 1.00 respectively. The small peaks, labeled "sidelobe peaks," appear around delay times for the first and third objects because the emitted frequency ranges only from 25 kHz to 106 kHz. The value of the cross-correlation function at the delay time for the second object is smaller than the sidelobe peak. This means that it is difficult to determine delay times for closely spaced objects with different reflectivity using that method.

Acknowledgments

This work was supported by KAKENHI (18700301). The author wishes to thank James A. Simmons, Hiroshi Riquimaroux, and Shizuko Hiryu for providing the bat sound data, as well as discussions.

References

[1] D.R. Griffin, *Listening in the dark* (Yale U. P., New Haven CT, 1958; reprinted by Cornell U. P., Ithaca, New York, 1986).

[2] J.A. Simmons, M.J. Ferragamo, P.A. Saillant, T. Haresign, J.M. Wotton, S.P. Dear, D.N. Lee, "Auditory dimensions of acoustic images in echolocation," in *Hearing by Bats*, edited by A.N. Popper and R.R. Fay (Springer, New York), pp. 146-190 (1995).

[3] J.A. Simmons, "Perception of echo phase information in bat sonar", *Science* 207, 1336-1338 (1979).

[4] J.A. Simmons, M.J. Ferragamo, C.F. Moss, S.B. Stevenson, R.A. Altes, "Discrimination of jittered sonar echoes by the echolocating bat, *Eptesicus fuscus*: the shape of target images in echolocation" , *J. Comp. Physiol. A* 167, 589-616 (1990).

[5] I. Matsuo, J. Tani, M. Yano, "A model of echolocation of multiple targets in 3D space from a single emission", *J. Acoust. Soc. Ame.* 110, 607-624 (2001)..

[6] P.A. Saillant, J.A. Simmons, S.P. Dear, T.A. McMullen, "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, 2691-2712 (1993).

[7] J.A. Simmons, M.J. Ferragamo, C.F. Moss, "Echo-delay resolution in sonar images of the big brown bat, *Eptesicus fuscus*", *Proc. Natl. Acad. Sci. U.S.A.* 95, 12647-12652 (1998).

[8] I. Matsuo, K. Kunugiyama, M. Yano, "An echolocation model for range discrimination of multiple closely spaced objects: Transformation of spectrogram into the reflected intensity distribution", *J. Acoust. Soc. Ame.* 115, 920-928 (2004).

[9] P.A. Saillant, J.A. Simmons, F.H. Bouffard, D.N. Lee, S.P. Dear, " Biosonar signals impinging on the target during interception by big brown bats, *Eptesicus fuscus*", *J. Acoust. Soc. Am.* 94, 2691-2712 (2007).