

**A Fundamental Study on Biomimetic Sensory Augmentation
Based on Bat Biosonar**

by

Miwa Sumiya

**Doshisha University
Graduate School of Life and Medical Sciences**

November 2017

Doshisha University
Faculty of Life and Medical Sciences

ABSTRACT

A Fundamental Study on Biomimetic Sensory Augmentation Based on Bat Biosonar

by Miwa Sumiya

Supervisor: Professor Yoshiaki Watanabe

Echolocating bats can “see the world by sound” using high and sophisticated ultrasonic sensing even in the absence of visual information. The final goal of this study is to give the sensation of “seeing by sound” not only to blind people but to all humans. In this dissertation, we conducted following three projects: 1) To learn from the bat about the possibility of sensing using ultrasound, we measured highly unique sonar behaviors of echolocating Japanese house bats (*Pipistrellus abramus*) during natural foraging using custom-made large-scale three-dimensional (3D) microphone array system and quantitatively analyzed them using mathematical model presented in this study. Our experimental and mathematical results suggested that echolocating bats coordinate their control of the acoustical field of view and flight for consecutive captures in 3D space during natural foraging. 2) Next, we constructed a psychoacoustic experiment system using ultrasonic binaural echoes measured by miniature dummy head (MDH). In this proposed psychoacoustic experiment, ultrasonic echoes from the target acquired by a 1/n-scaled MDH were presented after 1/n-times pitch conversion to listeners as binaural audible sounds through headphones. Using this system, we found that even sighted echolocation novices could discriminate visually different materials and surface structures different objects using ultrasonic binaural echoes measured by MDH. 3) Finally, based on the psychoacoustic experiment system, we examined if sighted echolocation novices could discriminate the 3D roundness of edge contours among five targets using the downward frequency-modulated (FM) ultrasound mimicking the echolocation sound of the bats. These targets were difficult to discriminate by vision without shadows. Our results showed that the participants could identify the roundness of edge contours by using downward FM echoes in the high-frequency range (35–7 kHz) that were converted to pitch in the audible range (5–1 kHz). In addition, the participants could discriminate between targets that were not used in the training. We additionally conducted shape, texture, and material discrimination experiments for other sighted echolocation novices to examine the suitable signal for the discriminations. Consequently, we found that the broadband signals were useful in shape, texture, and material discriminations, because timbre is a robust acoustic cue.

TABLE OF CONTENTS

List of Figures	V
List of Tables	VIII
Acknowledgements	IX
Chapter 1. Introduction	
1.1 Final goal of this study.....	1
1.2 Echolocation in bats.....	1
1.3 Motivation.....	1
1.4 Purpose.....	2
1.5 Organization.....	2
Chapter 2. Background	
2.1 Echolocation in bats.....	4
2.1.1 Echolocation when hunting insect prey.....	4
2.1.2 Field measurement of bats.....	5
2.1.3 Mathematical modeling of animal behavior.....	6
2.2 Echolocation in humans.....	6
2.2.1 Early research investigating human echolocation abilities.....	7
2.2.2 Human echolocation using mouth clicks.....	7
2.2.3 Human echolocation using artificial sounds except mouth clicks.....	9
2.2.4 Comparison of echolocation ability among bats, dolphins, and humans.....	10
2.2.5 Blind mobility aids using ultrasounds.....	11
Chapter 3. Coordinated Control of Acoustical Field of View and Flight in Three-dimensional Space for Consecutive Capture by Echolocating Bats during Natural Foraging	
3.1 Introduction.....	13
3.2 Materials and Methods.....	15
3.2.1 Subjects and study site.....	15
3.2.2 Large-scale microphone array system.....	17
3.2.3 Reconstruction of 3D flight paths and pulse directions.....	18
3.2.4 Acoustic analysis.....	20
3.2.5 Methods for the numerical simulation.....	20
3.3 Results.....	22
3.3.1 3D flight paths and pulse directions of bats attacking multiple targets in the field....	22
3.3.2 Future prey positions were covered by acoustical field of view.....	23
3.4 Discussion.....	25
3.4.1 Relationship between acoustic sensing and bat flight.....	25
3.4.2 Acoustic sensing in the horizontal and vertical planes.....	27
3.5 Summary.....	28

Chapter 4. Mathematical modeling of flight and acoustic dynamics of an echolocating bat during multiple-prey pursuit

4.1 Introduction.....	36
4.2 Materials and Methods.....	37
4.2.1 Mathematical modeling of flight and acoustic dynamics.....	37
4.2.2 Numerical simulation.....	39
4.2.2.1 Initial conditions.....	39
4.2.2.2 Conditions of prey capture.....	40
4.2.3 Field measurement.....	40
4.3 Results.....	41
4.3.1 Results of simulation.....	41
4.3.2 Flight behavior of the bat while approaching multiple targets.....	41
4.4 Discussion.....	41
4.5 Summary.....	43

Chapter 5. Human echolocation by ultrasonic binaural echoes

5.1 Introduction.....	48
5.2 Materials and Methods.....	50
5.2.1 Participants.....	50
5.2.2 Targets.....	50
5.2.3 Miniature dummy head.....	50
5.2.4 Sound stimuli.....	51
5.2.5 Procedure.....	51
5.2.6 Statistical analysis.....	52
5.3 Results.....	52
5.4 Discussion.....	53
5.4.1 Discrimination performance using ultrasonic binaural echoes	53
5.4.2 Comparison with the echolocation performance using mouth clicks.....	54
5.5 Summary.....	55

Chapter 6. Shape, Texture, and Material Discrimination by Ultrasonic Binaural Echoes in Sighted Echolocation Novices

6.1 Introduction.....	67
6.2 Materials and Methods.....	70
6.2.1 Participants.....	70
6.2.2 Targets.....	70
6.2.3 Miniature dummy head.....	71
6.2.4 Sound stimuli.....	72
6.2.5 Procedure.....	73
6.2.6 Statistics.....	74
6.3 Results.....	75
6.3.1 Discrimination of 3D edge contours by low- and high-frequency signals.....	75
6.3.2 Effective signal design for discriminations of shape, texture and material.....	76
6.3.3 Frequency analysis of echoes.....	77
6.3.4 Acoustic cue.....	77
6.4 Discussion.....	78
6.4.1 New possibility of human echolocation using high-frequency signal	78
6.4.2 Robust acoustic cue for shape, texture and material discriminations.....	79
6.4.3 Comparison with the bat's ultrasonic sensing.....	81
6.5 Summary.....	82

Chapter 7. Conclusions

7.1 Summary of Main Results.....	94
7.1.1 Coordinated control of the acoustical field of view and flight in three-dimensional space for consecutive capture by echolocating bats during natural foraging (Chapter 3).....	94
7.1.2 Mathematical modeling of flight and acoustic dynamics of an echolocating bat during multiple-prey pursuit (Chapter 4).....	95
7.1.3 Human echolocation by ultrasonic binaural echoes (Chapter 5).....	95
7.1.4 Shape, texture, and material discriminations by ultrasonic binaural echoes in sighted echolocation novices (Chapter 6).....	96
7.2 Future studies.....	97
7.2.1 Background, issue, and resolution studies.....	97
7.2.2 Research content.....	98
7.2.3 Features and originality of the research.....	99
7.3 Final remarks.....	100
References.....	101
Curriculum vitae.....	113

List of Figures

- Figure 3.1. Large-scale 3D microphone-array system.
- Figure 3.2. Typical example of multiple consecutive capture flight of *P. abramus* in the field.
- Figure 3.3. Relationship between target positions and bat sonar beam while approaching targets.
- Figure 3.4. Relationship between the prey direction relative to pulse direction and the directivity patterns of the sonar beam as bats converge on immediate prey in short- and long-interval captures.
- Figure 3.5. Numerical simulation of relationship between prey direction relative to pulse direction and directivity patterns of sonar beam as bats converge on immediate prey during failure and success cases.
- Figure 4.1. Schematic diagram of the present mathematical model presented in this study.
- Figure 4.2. Microphone array system in the field.
- Figure 4.3. Representative cases of the flight path and pulse direction calculated from the present model.
- Figure 4.4. Representative case of the flight path and pulse directions during successive captures of two targets.
- Figure 5.1. Photographs of six targets and an anechoic chamber.
- Figure 5.2. Schematic diagram of the experimental system.
- Figure 5.3. Average percent of correct answers performance of the eight participants under the downward, upward FM sweeps and CF burst signal conditions.
- Figure 5.4. Scatter plots of each relationship between the echoes from six different targets or one from no targets based on the nonmetric multidimensional scaling method.
- Figure 5.5. Normalized amplitude spectra of echoes.
- Figure 5.6. Scatter plots of relationships among seven echoes under three signal conditions.
- Figure 5.7. Spectrogram of click sound by a subject.
- Figure 5.8. Averaged rate of correct answers of sound discrimination for four subjects under down FM, up FM, and CF conditions (ultrasonic binaural echo condition) and mouth click condition.
- Figure 6.1. Schematic diagrams and a photograph of targets used in this study.
- Figure 6.2. Photograph of the 1/7 scaled MDH which were inserted two microphones into the ear canals.
- Figure 6.3. Amplitude wave form and spectrogram of the logarithmic sweep signal transmitted by the loud speaker.
- Figure 6.4. The schematic diagram of experimental system of the stereo and binaural recordings.
- Figure 6.5. Spectrograms of nine signals which were convoluted with an impulse response measured in advance.
- Figure 6.6. Average percentage of correct performance of the ten participants P1–10 in the low- and high-frequency signal conditions.

Figure 6.7. Frequency spectra the impulse response of Shapes 1–5 which were convoluted with the low- and high-frequency downward FM sweeps.

Figure 6.8. Average percentage of correct answers performance of the 30 participants (P11–40), under eight signal conditions in the shape, texture, and material discrimination experiments, using ultrasonic binaural echoes. Frequency spectra (energy spectral density) of the echo parts of the impulse responses of the Shapes 1–5, Textures 1–5 and Materials 1–4.

Figure 6.9. Average percentage of correct performance of the ten participants P11-20 in the natural and loudness conditions. Average percentage of correct performance of the participant P13 with absolute hearing under eight signal conditions in the shape discrimination experiment.

List of Tables

Table 3.1. Definitions of angular variables.

Table 5.1 Average percent of correct answers of the eight participants in each target pair under the downward FM, upward FM and CF signal conditions.

Table 5.2 Summary of GLMM statistics for the effects of the degree of similarity among the acoustic parameters on the number of correct answers in each target pair.

Table 6.1. Average percentage of correct answers in each target pair of the 10 participants under the high-frequency signal condition in Experiment 1.

Table 6.2. Summary of GLMM statistics for the effects of the degree of similarity among the acoustic parameters and the absolute value of the differences in the loudness on the number of correct answer in each target pair in the shape, texture, and material discrimination experiments.

ACKNOWLEDGEMENTS

I am most grateful to many people that have been support in my study, and I wish to thank them all. First, I would like to express my sincere gratitude to my supervisor, Professor Yoshiaki Watanabe for providing me this precious study opportunity as a doctoral student and for encouraging me with his great support throughout years, and Professor Shizuko Hiryu for her correct guidance, warm encouragements, and many discussions that make my research of achievement and my study life unforgettable. I would also like to thank Dr. Kaoru Ashihara for his shrewd advice on acoustic measurement and for encouraging me. I wish to thank Associate Professor Kohta I. Kobayasi for his valuable advices and constructive comments on my research. I wish to thank them for providing such a great scientific environment for my thesis work.

I would like to express my sincere gratitude to Dr. Yoshiki Nagatani for his constructive comments on my research. I wish to express my warm and sincere thanks to Dr. Ikkyu Aihara for his warm support and shrewd advice on the mathematical research, and Dr. Emyo Fujioka for his insightful comments on the field measurement and heartwarming encouragement.

Many thanks go to my collaborators on field measurement project; Mr. Kazuya Motoi and Mr. Masaru Kondo for their patient effort and great support on field measurement. Many thanks go to my collaborators on human echolocation project; Mr. Yuki Sarumaru for valuable discussion and warm encouragement, Mr. Taito Banda for insightful discussions, great support and constant encouragement, Ms. Reina Uchikawa and Ms. Ayana Kamada for great support on psychoacoustic experiment and valuable discussions, and Mr. Masaki Gogami and Mr. Kazuki Yoshino for their insightful discussions and great support. I wish to express my warm and sincere thanks to like to thank Dr. Yasufumi Yamada and Dr. Takafumi Furuyama for their shrewd advice and warm encouragements, Mr. Hidetaka Yashiro and Mr. Shota Murai for their discussion in all scientific fields and daily heartwarming encouragement, and Mr. Kazuma Hase, Ms. Sachi Itagaki and everybody at the laboratory of Neuroethology and Bioengineering Laboratory in Doshisha University for their support.

I would like to thank Mr. Yoshitsugu Sawa and Mr. Hiroshi Araki for their valuable comments from an engineering standpoint.

I would like to thank Japan Society for the Promotion of Science (JSPS) for grants they provided, which made it possible for me to complete my thesis..

Finally, I would like to thank my parents: Kazuyuki and Yuko for their warm supports.

Miwa Sumiya

Chapter 1. Introduction

1.1 Final goal of this study

Humans perceive the world by sensing the surrounding environment using five senses (visual, auditory, tactile, olfactory, and taste) and processing the acquired information in the brain. However, the world that we perceive is one that only we can understand; it can be said that other animals perceive their own world which only they can understand by using information that were acquired using their senses to survive. For example, snakes possess a unique sensory system for the detection of infrared radiation, which enables them to generate a thermal image of predators or prey (Gracheva *et al.*, 2010). Weakly electric fish can determine their surroundings at night by employing active electrolocation (von der Emde, 2006). Mimicking the sensing strategy of such animals and applying it to engineering fields such as robot engineering is called biomimetics. In this study, we aim to give a new sense to humans by applying the sensing strategy of animals to human rather than to engineering technology.

1.2 Echolocation in bats

Bats are terrestrial mammals like humans, but they can “see the world by sound” even in the absence of visual information. In bats, it is also called sonar, which is an acronym for sound navigation and ranging. Bats have a greater ability for ultrasonic sensing compared to our current technologies (Bates *et al.*, 2011). Bats understand their surroundings and acquire target information using echolocation (detecting targets by analyzing echoes from targets) using ultrasounds. Bats have a simple sonar system consisting of only one transmitter (i.e., nose), two receivers (i.e., ears), and tiny central processor (i.e., brain).

1.3 Motivation

Some blind people have independently acquired the ability for echolocation. Recently, the interest for human echolocation has increased. For example, videos of blind people using mouth clicks to play basketball or ride a bicycle have been uploaded on a video sharing site. Moreover, it was discovered that the visual cortex of blind people who

use echolocation was activated when they listened to the echo from an object (Thaler *et al.*, 2011), suggesting that the new sensation of “seeing the sound” by echolocation can be substituted for vision in the brain. Based on these facts, we study the possibility that humans can “see by sound.” We focus on the ability to “see by sound” using echolocation as the first step of this research and apply the sensing ability of the living beings to humans.

1.4 Purpose

We think that it is necessary to propose a sensing method and an efficient training method for a specific purpose (e.g., object localization, object identification) to give the sensation of “seeing by sound” not only to blind people but to all humans. The bat, which uses echolocation, uses ultrasounds with higher spatial resolution. However, we cannot listen to ultrasounds. Therefore, in this dissertation, we first constructed a psychoacoustic experiment system that allows humans to experience echolocation using ultrasound. Next, using the proposed psychoacoustic experiment system, we investigated how sighted people who have never performed echolocation have perceived the difference of objects with ultrasound. We analyzed the acoustic parameters of echo that are effective as acoustic cues in each task. In addition, we examined the time-frequency structure of the transmission signal that can maximize the use of the acoustic cues. Moreover, when we conducted research on the human subjects included in this study, we thought that it was essential to learn from the bat about the possibility of sensing using ultrasound. Thus, we measured highly unique sonar behaviors and quantitatively analyzed them using mathematical models.

1.5 Organization

This dissertation is organized as follows. Chapter 2 contains a brief description of echolocation in bats and humans and a review of related literature. Chapter 3 describes the coordinated control of the acoustical field of view and flight in three-dimensional space for consecutive capture by echolocating bats during natural foraging. Chapter 4 presents the mathematical models of flight and acoustic dynamics of an echolocating bat during multiple-prey pursuit. In Chapter 5, target discrimination performance using the ultrasonic binaural echoes by sighted echolocation novices were examined. Chapter 6

presents effective signal design for shape, texture, and material discrimination in human echolocation. Finally, conclusions and possible directions for future research are presented in Chapter 7.

Chapter 2. Background

2.1 Echolocation in bats

In this section, we introduce the basic sonar behavior of bats during natural foraging. In addition, review of field measurements of the bats sonar behavior that has been changing with the advancement of measurement technology was described. The mathematical biology and biosonar research in bats are also described.

2.1.1 Echolocation when hunting insect prey

Bats (*Chiroptera*) represent one of the largest and most diverse radiations of mammals, accounting for one-fifth of extant species (Simmons *et al.*, 2008). Bats can capture flying and moving small insects (i.e., their prey) by measuring the space using their advanced acoustic sensing ability even in the dark. The echolocation pulse of bats are categorized into two types: frequency modulation (FM) and constant frequency (CF) (Simmons *et al.*, 1979). Because the FM pulse consists of many frequency components, it can acquire more detailed information than the CF pulse. This suggests that the FM pulses are suitable for target ranging (Simmons, 1973). On the other hand, the CF pulses can concentrate energy because the frequency is stable; thus, it might be useful for the detection of a distant target. Japanese house bats (*Pipistrellus abramus*) have been reported to use a pulse with many CF components, which is called a quasi-CF pulse, during the search phase (Fujioka *et al.*, 2011).

The foraging behavior of bats was categorized into the following three phases based on the characteristics of the sonar sound: a search phase, an approach phase, and a terminal phase (Schnitzler and Kalko, 2001). Hiryu *et al.* (2008a) showed that similar to other pipistrelles (Schnitzler *et al.*, 1987), echolocation could be described with three particular patterns: search, approach, and terminal phases. In addition, *P. abramus* were reported to change the acoustic parameters (e.g., inter-pulse interval [IPI], pulse duration, terminal frequency) depending on the flight condition and target distance. This sensing strategy was not used in the existing artificial sonar system and it is interesting from an engineering perspective.

2.1.2 Field measurement of bats

To understand the biosonar behavior of bats and dolphins, it is important to measure the sonar sounds they emit accurately in addition to tracking their movements. There have been innovations in studies of biosonar and measurement technology for tracking animal movements as well as sound recordings. For example, three-dimensional (3D) flight trajectories of bats were reconstructed in a laboratory chamber using multiple video cameras (i.e., direct linear transform [DLT] method) (Tian and Schnitzler, 1997; Ghose and Moss, 2003; Hiryu *et al.*, 2005). In contrast, behavioral measurement for bats in the field is more difficult than in the laboratory because foraging wild bats fly in a large area at night. To overcome the difficulty of visual observation of nocturnal animals, Simmons (2005) employed infrared video cameras to monitor the movements of bats and large flying insects from their body heat. In addition, Schaub and H. U. Schnitzler (2007) reconstructed the 3D flight path of bats in the field using two infrared video cameras while recording the bat's broadcast.

Recently, microphone array technology has been used because it allows us to measure the 3D flight trajectory using the information of the bat's sonar broadcast. In particular, the bat's 3D flight trajectory can be reconstructed by calculating the position of a sound source (i.e., the position where the bat emitted the sonar sound) based on the time differences of sound arrival between the microphones. This was used to measure the flight trajectory of bats foraging in the field because it allowed us to measure the flight behavior even at night (Surlykke *et al.*, 2009b; Fujioka *et al.*, 2011; Seibert *et al.*, 2013). Additionally, the microphone array allows us to measure the directivity of bat's sonar beam. Laboratory experiments using a microphone array system revealed that the bats actively adjust the direction of their sonar beam to localize a target effectively (Ghose and Moss, 2006; Surlykke *et al.*, 2009a; Jakobsen and Surlykke, 2010; Yovel *et al.*, 2011; Matsuta *et al.*, 2013). We constructed a 32-channel microphone array covering a field area of 22 m × 24 m so that not only 3D flight paths but also horizontal pulse direction of bats during natural hunting could be measured. *P. abramus*, which are aerial-feeding bats, directed their horizontal pulse direction (i.e., acoustic attention) to the flight and other directions while searching for insect prey (Fujioka *et al.*, 2014). *P. abramus* have also been reported to capture successive flying insects every few seconds by dynamically and acrobatically changing their flight paths in open space. These experimental results suggest

that echolocating bats have ultrasonic sensing strategy for multiple targets.

2.1.3 Mathematical modeling of animal behavior

In recent years, biomimetics engineering has proposed to apply various mechanisms learned from animals to engineering. To apply the behavioral strategies of animals mechanically, it is essential to formulate them as algorithms. Collective motion of animal livings in groups, e.g., birds (Bajec and Heppner, 2009; Hildenbrandt *et al.*, 2010) and fish (Hemelrijk and Kunz, 2004), have been well studied in the field of mathematical biology (Vicsek and Zafeiris, 2012). However, there are few mathematical studies on the individual motion of animals.

Techniques for flexible acoustic sensing of bats are said to be very valuable in terms of engineering utilization. Mathematical modeling and quantitative analysis of sonar behaviors of bats have been performed. Ghose and Moss (2006) investigated the relationship between the pulse direction and the flight direction of a bat flying in the observation room using existing simple mathematical expressions that show the relationship between the pedestrian's line of sight and the direction of movement. However, there are only a few papers using such models about bat's sonar behavior. In addition, in laboratory experiments, it was not possible to maximize the performance of bats due to various factors such as narrowness of the observation room and acclimatization to the environment. To learn various animal strategies, such as control of acrobatic flight and pulse direction to capture prey efficiently, it is necessary to learn its significance from a wild bat not a captive bat. Thus, it is necessary to improve the techniques of measurements in the field and to perform quantitative analysis for data obtained by field measurement.

2.2 Echolocation in humans

As mentioned in Section 2.1, the unique and advanced sonar behavior of bats is very interesting even from an engineering perspective, and it has been studied through quantitative laboratory experiments and field measurements. Furthermore, to analyze the sonar behavior of bats more quantitatively, analysis using mathematical model has been conducted. There are blind individuals who have acquired high echolocation ability, like

bats. Some of them were reported to use echolocation on a daily basis, which enabled them to explore cities, hike, bike, or play basketball (Thaler *et al.*, 2011). Human echolocation is a hot topic in the field of psychology and neuroscience, and there have been several reviews of human echolocation (Kolarik *et al.*, 2014; Thaler, 2015; Thaler and Goodale, 2016). In this paper, we introduce the latest human echolocation research in addition to a review of previous human echolocation studies.

2.2.1 Early research investigating human echolocation abilities

Blind people have been reported to detect obstacles using “facial vision” until 1940. Some blind people who can detect obstacles without vision were considered to sense slight change in the air pressure by their skin, especially the skin of their face. However, experiments conducted by the research group at Cornell University in the 1940s and 1950s proved that “facial vision” was actually an “auditory ability” (Supa *et al.*, 1944; Worchel and Dallenbach, 1947; Cotzin and Dallenbach, 1950). About the same time that the mystery of facial skin was solved, echolocation was discovered by Griffin (1944). After that, it became clear that humans, especially the blind men, can use echolocation like bats. In the 1960s, Kellog and Rice conducted obstacle detection experiments using active human echolocation (Kellog, 1962; Rice and Feinstein, 1965; Rice, 1967; 1969).

2.2.2 Human echolocation using mouth clicks

To date, various studies have been performed for psychophysical examinations of acuity in perceiving location (distance and direction) (Teng and Whitney, 2011; Thaler *et al.*, 2011; Schornich *et al.*, 2012; Teng *et al.*, 2012; Wallmeier *et al.*, 2013; Wallmeier and Wiegrebe, 2014a; Wallmeier and Wiegrebe, 2014b; Pelegrín-García *et al.*, 2015), size (Rice and Feinstein, 1965; Teng and Whitney, 2011; Milne *et al.*, 2015a), shape (Rice, 1967; Arnott *et al.*, 2013; Milne *et al.*, 2014), and material (Hausfeld *et al.*, 1982; Milne *et al.*, 2015b) through various discrimination experiments using mouth clicks emitted by participants. Few studies were performed to study the echolocation ability of sighted people who have no experience with echolocation (Hausfeld *et al.*, 1982). However, most studies were conducted to measure the echolocation ability of blind expert echolocators who use echolocation in their daily lives; the echolocation performance of blind non-echolocators and sighted people were measured as controls to compare with blind expert

echolocators. In more recent studies, experiments were performed using virtual echo-acoustic space in which the participants gained spatial information about their environment by producing sounds with their mouths and evaluating computer-generated echoes of these sounds presented via headphones (Wallmeier *et al.*, 2013; Wallmeier and Wiegrebe, 2014b; Wallmeier and Wiegrebe, 2014a). In addition, some of the studies mentioned above performed both psychological experiments and fMRI measurements (Teng and Whitney, 2011; Thaler *et al.*, 2011; Milne *et al.*, 2015b).

The mouth clicks were reported to be useful signals for natural human echolocation (Kellog, 1962; Schenkman, 1986). However, these previous studies did not specify the kinds of clicks used in the experiments. Therefore, Rojas *et al.* (2009; 2010) conducted a physical analysis of organic sounds to determine which organic sounds were suitable for echolocation; they found that palatal clicks are the most convenient for natural human echolocation. Recently, the interest in acoustic characteristics of human expert echolocation sonar signals (i.e., mouth clicks) has been growing. The mouth clicks were reported to tend to be 3–15 ms long (Schornich *et al.*, 2012; Thaler and Castillo-Serrano, 2016) with peak frequencies ranging from 3 to 8 kHz (Thaler and Goodale, 2016). According to a large database of click emissions of three blind people expertly trained in echolocation and measured by Thaler *et al.* (2017), the mouth clicks produced by the human echolocation experts were reported to be consistently very brief (less than 3 ms duration) with peak frequencies of 2–4 kHz, but with energy at 10 kHz.

As mentioned above, Thaler *et al.* (2011) conducted fMRI measurements and demonstrated that blind echolocation experts use regions of the primary visual cortex when making echo-based judgments about object identity, suggesting that echolocation is a substitute for vision. Furthermore, some recent studies demonstrated that echolocation may substitute for vision on the behavioral level (Buckingham *et al.*, 2015; Milne *et al.*, 2015a). Milne *et al.* (2015a) showed that blind expert echolocators consistently identified the true physical size of the objects independent of distance, suggesting that echolocation exhibits “size constancy”. In addition, Buckingham *et al.* (2015) showed that blind expert echolocators were reported to experience a “size–weight illusion” when they used echolocation to get a sense of how big objects were and then judged their weight.

2.2.3 Human echolocation using artificial sounds except mouth clicks

Natural echolocation has several clear advantages: it does not need batteries, it is cheap, it cannot be forgotten at home, it does not break—and importantly, it can be learned by children (Thaler and Goodale, 2016). Moreover, mouth clicks do not interfere with other activities (i.e., they can be modulated and even stopped when the person wants to talk or do something else) (Thaler and Goodale, 2016). Therefore, as mentioned in section 2.2.2, measurement of natural human echolocation ability using mouth clicks and acoustical analysis of mouth clicks emitted by blind echolocation expert individuals have been conducted. In contrast, echolocation using artificial sound has also several advantages. Rojas *et al.*, who performed acoustical analysis of the mouth clicks (Rojas *et al.*, 2009), described the importance of echolocation using mouth clicks and discussed the advantages of echolocation using artificial sound (e.g., excellent reproducibility). There are some recent studies that examined echolocation performance, especially localization performance, using audible artificial sound (Arias and Ramos, 1997; DeLong *et al.*, 2007b; Schenkman and Nilsson, 2010; 2011; Rowan *et al.*, 2013; Rowan *et al.*, 2015; Schenkman *et al.*, 2016). Discrimination experiments using several artificial sounds allow us to examine acoustic cues quantitatively. These previous studies examined effective acoustic cues for detection (Arias and Ramos, 1997), localization, and material discrimination (DeLong *et al.*, 2007b). In addition, some studies have suggested that band-limited noise is more effective for detection and localization of targets compared to clicks. Furthermore, more high-frequency band-limited noise is effective (Rowan *et al.*, 2013; Rowan *et al.*, 2015). However, the band-limited noise used in these studies was within the audible range, suggesting that a higher frequency signal is necessary to improve echolocation performance. A few biosonar researchers who study dolphins have performed material and shape discrimination experiments for humans with dolphin-like signals in the ultrasonic range (Au and Martin, 1989; DeLong *et al.*, 2007b). However, this is rare in the field of human echolocation research. In addition, some biosonar researchers who study bats have also been studying the human echolocation, but their research is mainly focused on natural human echolocation using mouth clicks (Wallmeier *et al.*, 2013; Wallmeier and Wiegrebe, 2014b; Wallmeier and Wiegrebe, 2014a).

2.2.4 Comparison of echolocation ability among bats, dolphins, and humans

There have been many studies on the echolocation ability of bats and dolphins through various discrimination experiments. The estimated target range from the time delay between the outgoing vocalization and returning echo were measured. Four species of echolocating bats (*Eptesicus fuscus*, *Phyllostomus hastans*, *Pteronotus suapurensis*, and *Rhinolophus ferrumequinum*) were reported to be able to discriminate range differences as small as 1 to 3 cm (Simmons, 1973). In particular, echolocating *E. fuscus* have the ability to detect echo delay with a time resolution of 500 ns (Simmons, 1979). With respect to azimuth, the trained big brown bat exhibited an acuity of 1.5° (75% performance threshold) for discriminating horizontal angles separating rods in either two-rod or five-rod arrays (Simmons *et al.*, 1983). In addition, echolocating *R. ferrumequinum* detect insects by concentrating on the characteristic amplitude and frequency modulation (von der Emde and Schnitzler, 1990). A recent study examined the behavioral strategies for texture discrimination by echolocation in free-flying *E. fuscus*; they found that the flying bat listens to changes in sound spectra from an echo to another echo to discriminate between objects (Falk *et al.*, 2011). Another recent study performed numerical simulation and suggested that bats can discriminate between different tree types (Yovel *et al.*, 2008). Echolocating Atlantic bottlenose porpoises (*Tursiops truncatus*) can reliably discriminate between cylinders with different wall thicknesses or different materials in the water (Hammer Jr and Au, 1980).

On the other hand, blind people were reported to detect differences of approximately 10 cm in depth at a distance of 60 cm (Kellog, 1962). With respect to azimuth, Thaler *et al.* (2011) found that the blind expert echolocator can reliably distinguish a 3° difference in the position of the test pole straight ahead, even when listening only to recordings of echolocation sounds. In addition, Teng *et al.* (2012) found that experts were able to discriminate horizontal offsets of stimuli as small as 1.2° auditory angle in the frontomedial plane, which is comparable to the discrimination ability of horizontal angles measured by Simmons *et al.* (1983). With respect to shape discrimination, Milne *et al.* (2015a) found that a blind expert echolocator consistently identified the true physical size of the objects independent of distance, suggesting that echolocation exhibits “size constancy”. Additionally, Milne *et al.* (2015b) performed fMRI measurements while three blind echolocation experts identified a material (fleece,

synthetic foliage, or whiteboard) or an empty anechoic room; they found that the identification performance of participants was significantly above chance.

2.2.5 Blind mobility aids using ultrasounds

Various technological assistive devices for blind individuals have been developed based on echolocation for a long time (Kay, 1964; Kay, 1974; Ifukube *et al.*, 1991; Kay, 2000; Hughes, 2001; Mihajlik *et al.*, 2001; Waters and Abulula, 2007; Sohl-Dickstein *et al.*, 2015). These devices can emit ultrasounds and receive echoes from targets, which is similar to that of bats and dolphins; this allows us to localize targets with a high degree of accuracy by listening to pitch-converted ultrasonic echo.

These technological assistive devices can emit various ultrasounds with different time-frequency structures. For example, a device developed by Ifukube *et al.* (1991) transmitted upswept or downswept FM ultrasounds in the range of 70–40 kHz or a constant frequency (CF) sound of 50 kHz for echolocation. The signals that were acquired by stereo recording (because two receivers were not positioned at the ear) were time-stretched prior to presenting them to a user. Based on the results of their experiments, they found that downswept FM ultrasound was superior for the recognition of small obstacles compared to other ultrasonic schemes. By comparing the echolocation performance using various ultrasounds with different time-frequency structures, it is possible to investigate the optimum transmission signal suitable for each task.

To present pitch-converted ultrasonic echo to participants in a more realistic state, binaural recording of ultrasonic echoes is essential. Recently, Sohl-Dickstein *et al.* (2015) developed a device, referred to as the “Sonic Eye,” that uses a forehead-mounted speaker to emit ultrasonic logarithmic upward FM sweeps in the range of 25 to 50 kHz. The echoes are recorded by bilaterally mounted ultrasonic microphones, each mounted inside an artificial pinna, which were modeled after bat pinna. They demonstrated that the device can be effectively used to examine the environment and that the human auditory system rapidly adapts to these artificial echolocation cues.

In a previous study, we developed a new system for binaural recording in the ultrasonic range (Uchibori *et al.*, 2015). The system consists of a miniature dummy head (MDH) that is printed based on 3D shape data of a standard dummy head and a device that converts ultrasonic echoes into audible sounds (Uchibori *et al.*, 2015). The results

indicated that the sounds captured by the proposed MDH system were more accurately localized by the participants and outside the head more than normal stereo sounds (Uchibori *et al.*, 2015). Therefore, we concluded that MDH system may be useful for echolocation using ultrasound for humans.

Chapter 3. Coordinated Control of Acoustical Field of View and Flight in Three-dimensional Space for Consecutive Capture by Echolocating Bats during Natural Foraging

(Sumiya *et al.*, 2017)

3.1 Introduction

Echolocating bats have remarkable ultrasonic sensing abilities. They emit directional ultrasonic signals and listen to the echoes returning from objects. Perception by the bats firstly depends on the acoustical features of the returning echo, which change depending on various factors, such as distance to target, direction of target, and target strength. Second, the direction and directivity patterns of the beams emitted by these bats are simple but important factors that affect their “field of view” from echolocation (acoustical field of view), and restrict the extent of spatial information during echolocation. The microphone-array technique has been widely used to track the 3D flight trajectory of bats (Chiu *et al.*, 2010; Brinkløv *et al.*, 2011; Fujioka *et al.*, 2011; Seibert *et al.*, 2013) and also allows us to measure the direction and directivity of the bats’ sonar beams during flight in a flight chamber and in the field (Ghose and Moss, 2003; 2006; Ghose *et al.*, 2009; Surlykke *et al.*, 2009a; Surlykke *et al.*, 2009b; Chiu *et al.*, 2010; Jakobsen and Surlykke, 2010; Yovel *et al.*, 2011; Matsuta *et al.*, 2013; Seibert *et al.*, 2013). By measuring the pulse direction (*acoustic attention*) and directivity patterns (beam width) as useful indices, we can investigate where the bats direct their attention and potentially detect prey, and how they control their acoustical field of view during advanced flight maneuvers. Recently, it has been reported that echolocating bats actively change the pulse direction depending on the situation in a flight chamber (Ghose and Moss, 2006; Ghose *et al.*, 2009; Surlykke *et al.*, 2009a; Chiu *et al.*, 2010; Yovel *et al.*, 2011) or even in the field (Seibert *et al.*, 2013; Fujioka *et al.*, 2014). They have also been reported to expand the beam width adaptively before capturing target prey to retain a moving target within the acoustical field of view (Jakobsen and Surlykke, 2010; Matsuta *et al.*, 2013) or to narrow the beam width when entering a confined space (Kounitsky *et al.*, 2015). These studies demonstrated that bats actively adjust both the pulse direction and the beam width, as well as traditional acoustic characteristics, such as time-frequency

structure, sound pressure level, or inter-pulse interval (IPI) (Surlykke and Moss, 2000; Hiryu *et al.*, 2008b).

In our previous study, we constructed a 32-ch microphone-array covering an area of 22×24 m, to measure the 3D flight paths and horizontal pulse directions of the Japanese house bat, *P. abramus* during natural foraging (Fujioka *et al.*, 2014). *P. abramus* exhibits acrobatic hunting behavior, which involves occasionally capturing multiple airborne insects within a short time interval (less than 1 s) in the field (Fujioka *et al.*, 2011; Fujioka *et al.*, 2014). We also reported that the bats directed their pulse toward the subsequent target before capturing the immediate one when attacking two successive targets (Fujioka *et al.*, 2014). On the other hand, our recent study adopted a mathematical methodology to estimate parameters representing the *flight attention* of bats for their measured flight paths during the phase of approaching prey (Fujioka *et al.*, 2016). (Note that the *flight attention* is derived from a parameter of weighing factors to minimize the angular difference between the bat's flight direction and the direction to its prey. This represents the attention by the bat toward a certain target prey in terms of flight.) We found that the distribution of the *flight attention* parameters (estimated from the behavioral data of wild *P. abramus* during two consecutive captures in a short-time interval) corresponded to the optimal value of the parameter set in the numerical simulation, which showed a high success rate of consecutive prey captures. This indicated that bats plan their future flight paths based on additional information about their next prey. This model was based on the assumption that bats distribute their attention across multiple prey items. This implies that bats acoustically recognize the positions of multiple prey items at the same time, even while approaching their immediate prey. To confirm this assumption experimentally, we investigated the relationship between the prey direction (direction of prey relative to the pulse direction) and the beam width of the bat during natural foraging in the field. We proposed two hypotheses: 1) bats physically shift their pulse direction between immediate and subsequent targets in somewhat of a time-sharing manner (time-sharing manner hypothesis), or 2) bats aim their beam across a wide area to cover both immediate and subsequent targets simultaneously within their acoustical field of view (acoustical field hypothesis). Based on our previous studies (Fujioka *et al.*, 2014; Fujioka *et al.*, 2016), we propose an acoustical field hypothesis in which the bats are assumed to aim their pulses in a certain direction. This keeps multiple

targets within their acoustical field of view, including the peripheral part. These features may be linked to processes in visual-guided animals; for example, humans sometimes detect objects using the peripheral field of vision (Williams and Davids, 1998; Crundall *et al.*, 1999; Ryu *et al.*, 2015). Visual-guided animals are also known to shift the gaze direction sequentially to guide movement planning (Eckmeier *et al.*, 2008). Therefore, the investigation of control of the acoustical field of view by bats during flight not only provides insight into how animals actively sample spatial information from their environment during locomotion, but can also be used for comparative studies along with visual research on how animals use their vision during locomotion (Ghose and Moss, 2006; Surlykke *et al.*, 2009b).

Echolocating bats seem to use their acoustical field of view effectively during aerial-feeding flights by advanced coordinated control of the acoustical field of view and flight. To investigate how bats control these features in 3D space during consecutive prey-capture flights, we established horizontal and vertical microphone-arrays in the field to measure the flight paths and 3D pulse direction and directivity pattern of the emitted sounds during natural foraging. For every pulse emission, we examined time variation in the pulse direction and beam pattern in relation to the direction of immediate and subsequent targets while attacking multiple target prey consecutively in the field. Furthermore, we conducted a numerical simulation to determine how the bats control the acoustical field of view according to the prey directions for the successful capture of both immediate and subsequent prey items.

3.2 Materials and Methods

3.2.1 Subjects and study site

The subject of this study was *P. abramus*, which is a member of the family Vespertilionidae and has a wingspan of 10–15 cm and a bodyweight of 5–8 g. During natural foraging, *P. abramus* emits relatively long (9–11 ms), shallow-swept frequency-modulated (FM) pulses, with the energy concentrated in the terminal frequency of the fundamental component at around 40 kHz (Hiryu *et al.*, 2008a; Fujioka *et al.*, 2011). The bats are regularly observed in large open areas during the evening from early summer to fall. Here, the study site was a large open area over a river with the width of approximately

20 m, near the campus of Doshisha University in southern Kyoto Prefecture, Japan, where only *P. abramus* regularly appears to forage for airborne insects. Our field studies did not involve endangered or protected species. No specific permissions were required for these locations/activities because the study site was not protected by the regulatory body concerned with the protection of wildlife. The target prey of *P. abramus* are mainly small hemipterans and dipterans (Hirai and Kimura, 2004). The flight speed of the common chironomid midge *Chironomus plumosus* ranges from 0.25 to 1.1 m/s (Crompton *et al.*, 2003), and we visually observed that swarming dipteran midges took several seconds to fly across an area of a few tens of centimeters. From our observation, we concluded that these prey are roughly ten times slower than *P. abramus* (which has an average speed of 5 m/s) and therefore assumed that the movement of the prey, at this study site, was negligible during the brief period of approach and capture (< 3 s) in each flight sequence examined in this study (Fujioka *et al.*, 2016).

When a bat successfully captures an insect, a brief burst of sounds (feeding buzz) occurs, followed by a silent interval (post-buzz pause) (Schnitzler *et al.*, 1987; Kalko and Schnitzler, 1989; Britton and Jones, 1999). Based on our previous recording, *P. abramus* emits a feeding buzz approximately 0.2 s before capture (Fujioka *et al.*, 2011; Fujioka *et al.*, 2014), at a distance of a few tens of centimeters (approximately 30 cm) from the prey point. Since the movement of prey can be negligible within this brief period of time, we simply defined the bat's 3D position at the end of the feeding buzz as the location of the prey at the capture in this study (Fujioka *et al.*, 2011; Fujioka *et al.*, 2014; Fujioka *et al.*, 2016).

It is difficult to draw definitive conclusions about whether a bat has successfully captured prey based only on information about the post-buzz pause (Britton and Jones, 1999; Fujioka *et al.*, 2016). However, we confirmed that the distances between successful prey positions were too large for the prey (small midges) to move during the observed inter-capture intervals for the flights analyzed in this study. Therefore, if the bats failed to capture prey, it is highly unlikely that they attacked the same prey again in the second capture trial just after the first unsuccessful capture. Although it is still an assumption, we assumed that the bats attacked different prey sequentially, at least in the case of two successive target captures with short time intervals (< 1.5 s) in this study. We thus categorized two successive target captures with long time intervals (> 3.0 s) as *long-*

interval capture and those with short time intervals (< 1.5 s) as *short-interval capture*.

3.2.2 Large-scale 3D microphone-array system

The recordings were carried out on 5 separate days (October 14, 2013; July 15, September 30, October 7, and October 16, 2014) for approximately one hour before and after sunset. Recording was conducted for 5–10 min each session and the total number of recording sessions was 32 for 5 days (total recording time: 225 min). To reconstruct the flight paths of the bats and identify the capture points accurately, we selected prey-capture flights where the amplitude of the sonar sounds was sufficiently high with a good signal-to-noise ratio, even at the end of the terminal buzz. At the same time, to ensure measurement accuracy, we selected only flights where the bat approached the targets while flying toward the L-shaped array inside of the U-shaped array. Thus, the vertical and horizontal pulse directions and beam widths could be measured accurately. We analyzed a total of 2,680 pulses from 37 captures in 20 measured flight paths. That is, flight paths with three successive captures were split into two flight paths with two successive captures each (See Fig 3.2.).

Echolocation pulses emitted from the bats were recorded using a custom-built 44-ch 3D microphone-array (Fig 3.1A). We previously reconstructed the horizontal pulse direction along with 3D flight paths using a horizontal U-shaped 32-ch microphone-array system (Fujioka *et al.*, 2014). In this study, we newly built a vertical 12-ch L-shaped microphone-array unit so that the vertical pulse direction could be measured. The data shown here are from measurements using two different configurations with 32 microphones in 2013 (vertical L-shaped 10-ch and horizontal U-shaped 22-ch), and 44 microphones in 2014 (vertical L-shaped 12-ch and horizontal U-shaped 32-ch).

The microphone-array units were arranged to cover the entire foraging area over the stream, as shown in Fig 3.1A. We reconstructed the 3D flight paths using four Y-shaped array units using omnidirectional electret condenser microphones (models FG-23329-C05 and FG-23629-P16; Knowles, Itasca, IL, USA). The emitted echolocation pulses were recorded and amplified using a custom-built electronic circuit via a 10–250 kHz band-pass filter, and were digitized with 16-bit accuracy at a sampling rate of 500 kHz using high-speed data acquisition cards (PXIe-6358; National Instruments, Tokyo, Japan). The output signals were synchronously stored using a personal computer via a

custom program using LabVIEW 2011 (National Instruments).

3.2.3 Reconstruction of 3D flight paths and pulse directions

The 3D locations of the bats were obtained using the difference in arrival times between a central and three other microphones separated by 1.3 ± 0.01 m in the Y-shaped array (Fig 3.1B). The arrival-time differences were calculated from cross-correlation functions using a Matlab routine (Math Works, Natick, MA, USA), and the analytical procedure was the same as in our previous studies (Fujioka *et al.*, 2011; Fujioka *et al.*, 2014; Fujioka *et al.*, 2016). We combined 3D sound coordinates calculated by each of four Y-shaped array units so that the flight trajectory could be reconstructed within the foraging area, which was enclosed by the microphone-arrays. The maximum range error of the microphone-array system was less than 10 cm for sound sources within 5 m of the Y-shaped array unit (Fujioka *et al.*, 2011).

The horizontal pulse direction was calculated based on the sound pressure difference, namely, the differences in peak power in the spectrogram of a sonar sound across 24 microphones that were distributed over the entire array at the same horizontal level. The vertical pulse direction was also measured by the L-shaped array unit by the same method (Fig 3.1C). The sound pressure levels of the pulses were corrected for the propagation loss of sounds in the air between the bat and each microphone and the sensitivity differences between the microphones in the array. Sound absorption was calculated from measured absorption coefficients that were determined for the average frequencies at the peak energy in the FM pulse of *P. abramus* (1.2 dB/m at 45 kHz). The sensitivity of the microphone-array elements was measured by sending a 10-ms burst at 45 kHz to each microphone of the array using an ultrasonic loudspeaker (PT-R7, Pioneer, Tokyo, Japan). This allowed the recorded sounds to be calibrated according to sensitivity differences between the microphones. The directivity pattern of the sonar beam was reconstructed by curve fitting using a Gaussian function, and the direction at the peak value of the reconstructed directivity pattern was determined as the pulse direction. The beam width was defined by -6 dB off-axis angles in the reconstructed directivity pattern from the pulse direction. In our previous studies, we calibrated the measurement accuracy of the pulse direction and beam width for a U-shaped array. The error of the pulse direction was less than $\pm 5^\circ$, and that of the beam width was less than $\pm 7^\circ$, in the horizontal

plane when the sonar beam was directed toward the U-shaped horizontal microphone-array (Fujioka *et al.*, 2014; Fujioka *et al.*, 2016). In the present study, prior to the experiments, we arranged the L-shaped arrays in the field following the actual recording setup to measure the error of the vertical pulse direction and the beam width by using artificial FM sounds emitted from a loudspeaker (PT-R7; Pioneer, Tokyo, Japan). As a result, when the bat's sonar beam was directed to the inside of the L-shaped array (as shown in Fig 3.1C), the errors for both the vertical pulse direction and the beam width were less than $\pm 5^\circ$. The amount of error in this study corresponded to that of our previous studies (Fujioka *et al.*, 2014; Fujioka *et al.*, 2016), which was an acceptable amount of error to investigate the acoustical behavior of bats at the study site. To ensure measurement accuracy, we measured the horizontal beam width only when the vertical pulse direction was within $\pm 45^\circ$ from the horizontal plane of the U-shaped horizontal microphone-array. Therefore, when the bats emitted a pulse toward the exterior of the microphone-array (i.e., the positive direction in the X -axis in Fig 3.1A), we did not use the data for calculating beam width, but rather used only that for pulse direction to ensure the measurement accuracy of the beam width because it was unclear whether the vertical pulse direction was directed within $\pm 45^\circ$ from the horizontal plane. The vertical pulse direction was reconstructed only when the horizontal pulse direction was toward the inside of the U-shaped microphone-array. The vertical beam width was also calculated only when the horizontal pulses were directed within $\pm 45^\circ$ from negative direction of X -axis (Fig 3.1A) to ensure measurement accuracy.

In a strict sense, because auditory perception depends on the level of the returning echo and the hearing threshold of the bat (Kick, 1982), we should consider not only the beam width but also the angular range, combined with the transfer function of the outer ear, to define the acoustical field of view (Wotton *et al.*, 1997; Aytekin *et al.*, 2004). In this study, however, we adopted the -6 dB point of the beam width as a simple index of the acoustical field of view during echolocation following the approach used in recent studies (Jakobsen *et al.*, 2013; Jakobsen *et al.*, 2015).

We defined the positional relationship between the bat and the prey position, as shown in Fig 3.1D. The variables φ and θ represent the horizontal and vertical angles, respectively. The bat's horizontal (vertical) gaze angle φ_{gaze} (θ_{gaze}) was defined as the pulse direction relative to the flight direction. When the bat captured multiple successive

targets, we used suffixes to represent the order of the captures. φ_{fp} (or θ_{fp}) was defined as the direction of the prey position relative to the flight direction of the bat, while φ_{pp} (or θ_{pp}) indicates the direction of the prey position relative to the pulse direction. (See Table 3.1.)

3.2.4 Acoustic analysis

The analytical procedure was also the same as that used in our previous studies (Fujioka *et al.*, 2011; Fujioka *et al.*, 2014). The acoustic characteristics, including IPI and pulse duration, were analyzed from echolocation sounds recorded by the central microphone of the Y-shaped array unit that was spatially closest to the sound source (the bat) and thus received the strongest version of each broadcast using an additional custom program in Matlab. IPI was determined as the time between the amplitude envelope peaks of successively emitted echolocation pulses on oscillograms. The pulse duration was determined from a spectrogram of the first harmonic component of an extracted individual pulse (1024 points fast Fourier transform with a Hanning window, 97% overlap) at -25 dB relative to the peak intensity of the pulse. The foraging behavior of the bats was categorized into three phases based on the characteristics of the sonar sound; namely, a search phase, an approach phase, and a terminal phase (Schnitzler and Kalko, 2001). The terminal phase of pipistrelles is usually divided into two signals: buzz I and buzz II (Kalko and Schnitzler, 1989). Buzz II signals have lower bandwidth and frequency than buzz I signals (Schnitzler and Kalko, 2001). Since the start of the approach phase of wild *P. abramus* was characterized by the appearance of a slight increase in the pulse duration (Fujioka *et al.*, 2011; Fujioka *et al.*, 2014), we used this as a simple index for identifying the timing of the beginning of the approach phase in this study.

For statistical comparisons, a *t*-test and Mardia–Watson–Wheeler test were used, when appropriate, to test for significant differences in beam width and angular variables between data sets.

3.2.5 Methods for the numerical simulation

We numerically simulated the flight trajectory of a bat when approaching two prey items successively, so that the relationships between the acoustical field of view and the direction of each prey could be investigated quantitatively. In this framework, the bat

dynamically changes its flight direction depending on the directions of the two prey items, using the mathematical model proposed in our previous study (Fujioka *et al.*, 2016). Briefly, the modeled flight dynamics in the horizontal and vertical planes can be described as follows:

$$\frac{d\varphi_b(t)}{dt} = \frac{1}{\delta_h} \{ \alpha_h \sin[\varphi_{bp1}(t) - \varphi_b(t)] + \beta_h \sin[\varphi_{bp2}(t) - \varphi_b(t)] \}, \quad (1)$$

$$\frac{d\theta_b(t)}{dt} = \frac{1}{\delta_v} \{ \alpha_v \sin[\theta_{bp1}(t) - \theta_b(t)] + \beta_v \sin[\theta_{bp2}(t) - \theta_b(t)] \}, \quad (2)$$

where δ represents positive weighting factors and α is the minimization of the angular difference between the bat's own flight direction $[\varphi_b(t), \theta_b(t)]$ and the direction to prey 1 $[\varphi_{bp1}(t), \theta_{bp1}(t)]$ (similar for β to prey 2). To simplify the numerical simulation, we constrain the parameters as follows:

$$\alpha_h^2 + \beta_h^2 = 1, \quad (3)$$

$$\alpha_v^2 + \beta_v^2 = 1. \quad (4)$$

Therefore, parameters α_h and β_h (α_v and β_v) are described as $\alpha_h = \sin\gamma_h$ and $\beta_h = \cos\gamma_h$ ($\alpha_v = \sin\gamma_v$ and $\beta_v = \cos\gamma_v$). To examine the ratio of the *flight attention* between the two prey items, the arctangent of parameters α and β is defined by γ : namely, γ_h and γ_v represent the arctangents in the horizontal and vertical planes, respectively.

The flight path was numerically simulated for a situation in which two prey items are distributed inside the bat's 3D sonar beam, which was modeled as a circular piston oscillating in an infinite baffle (Jakobsen and Surlykke, 2010). A simulation trial began when the bat started its approach phase to capture prey 1. The calculation conditions for a parameter set (γ_h and γ_v) are the same as those in our previous study (Fujioka *et al.*, 2016), and the beam widths of the sonar beam in the horizontal and vertical planes were defined as $\pm 59^\circ$ and $\pm 25^\circ$, respectively, which were derived from the experimental results of this study. Three trials were performed for each parameter set of γ_h and γ_v , ranging from $-\pi$ to π , respectively, with 0.01π steps, and 201×201 pairs. All variables in this model were calculated by using the fourth-order Runge–Kutta method, with a time step of 0.01 s (the flight speed of the model bat was taken to be 5 m/s based on the experimental data). Pulses to obtain the target positions were emitted at every step. We examined the relationships between the acoustical field of view and the direction of each prey, in cases

of both success and failure of the prey-capture trials, using this mathematical model. A parameter set was defined as a success when the bat captured (close in within 10 cm) the immediate prey (prey 1) and then the subsequent prey (prey 2) in sequence (capturing both prey items), without losing the location of the prey which the bat intend to capture (i.e., prey 1 before the capture of prey 1 and prey 2 after the capture of prey 1). A simulation trial was defined as a failure when prey 2 was located outside the sonar beam after the capture of prey 1 (capturing only prey 1) (note that cases in which both prey 1 and prey 2 were missed were not considered in the investigation). We categorized the numerical simulation results of all trials into cases of success and failure and then analyzed the relationships between the acoustical field of view and prey directions while approaching prey 1 and prey 2 sequentially. Initial positions of prey 1 and prey 2 were randomly determined in horizontal and vertical space in every trial.

3.3 Results

3.3.1 3D flight paths and pulse directions of bats attacking multiple targets in the field

Figures 3.2A–C show representative data of the 3D flight paths with pulse directions of *P. abramus* when the bat captured four insects (Targets 1–4) consecutively in 12 s. The pulse directions did not always coincide with the flight direction, especially in the horizontal plane. The bat changed its flight direction and pulse direction dynamically on either the right or the left side from its flight direction or shifted between different directions. On the other hand, 82% of all pulses whose vertical pulse direction could be measured in this flight case (shown as a blue arrow in Figs 3.2A–C) were emitted downward ($< 0^\circ$), whereas only 18% of the pulses were emitted upward ($\geq 0^\circ$). This suggests that the bat emitted most of its pulses downward from the horizontal plane in the vertical plane.

Figure 3.2D shows time series data of IPI during the flight shown in Figs 3.2A–C. When the bat started the approach to capture an insect, the IPI was decreased from approximately 100 to 5 ms. In this flight case, the time intervals between successive captures (Captures 2–3 and 3–4) were both 1.1 s, whereas the time interval for the first two successive captures (Captures 1–2) was 5.4 s. Throughout this flight period, the horizontal gaze angle, φ_{gaze} , was widely distributed from approximately -130° to 80° (Fig

3.2E). In contrast, in the vertical plane, the gaze angle θ_{gaze} had a narrower range from about -60° to about 40° (Fig 3.2F), indicating that the bat changed the horizontal acoustical field of view more dynamically than the vertical one during foraging.

Figures 3.2E and F show that φ_{gaze} and θ_{gaze} often corresponded to the directions of some of the subsequent prey positions; that is, not only the prey position of the immediate target but also subsequent target prey positions corresponded to the center of the acoustical field of view. In addition, both the immediate and subsequent prey positions were simultaneously within the -6 dB beam width of the sonar beam (gray vertical lines). In particular, Fig 3.2E shows that the pulse direction corresponded to the positions of Captures 3 and 4, approximately 4 s before Capture 2 occurred. (See from 1.5 s to 3.5 s on the horizontal axis.)

3.3.2 Future prey positions were covered by acoustical field of view

Figures 3.3A–C show time series data of target directions before capture of the immediate target (0.5 s) and we could define the capture points of subsequent targets as the positions of the prey themselves in the case of *short-interval capture*. Figure 3.3A shows that in the case of *long-interval capture*, the subsequent prey (φ_{pp2} and θ_{pp2}) were outside the acoustical field of view during the approach period of the immediate target. This suggests that the bat acoustically focused only on the immediate prey during *long-interval capture*. On the other hand, during *short-interval capture*, the subsequent target prey positions were within the beam width of the bat's emissions during the whole or part of this period in both horizontal and vertical planes (Figs 3.3B and C). These results suggest that the bats can detect the positions of both their immediate and subsequent ones at the same time by keeping them within their acoustical field of view before capturing the immediate one during a *short-interval capture*. Figure 3.3A also shows that the immediate target directions (φ_{pp1} and θ_{pp1}) were not always maintained at the center of the acoustical field of view, but were within the bat's -6 dB beam width, in both horizontal and vertical planes, in the case of *long-interval capture*.

Figures 3.4A–D show the relationship between the target direction and -6 dB beam width during the approach phase and buzz I in terminal phase for all recorded flight data (20 flights). The time of the measured flight paths ranged from 2 to 18 s, with an average of 7.5 s. In the *long-interval captures* (10 flights), the horizontal and vertical

mean -6 dB beam widths of the pulses were $\pm 51 \pm 13^\circ$ ($N = 150$ pulses) and $\pm 27 \pm 14^\circ$ ($N = 95$ pulses), respectively. On the other hand, in the *short-interval captures* (10 flights), the horizontal and vertical mean -6 dB beam widths of the pulses were $\pm 59 \pm 12^\circ$ ($N = 131$ pulses) and $\pm 25 \pm 8^\circ$ ($N = 85$ pulses), respectively. Namely, the horizontal mean beam widths were significantly wider than the vertical ones for both *short-* (the t -test, $t(214) = 25.1$, $p < 0.001$) and *long-interval capture* cases (the t -test, $t(195) = 13.7$, $p < 0.001$). In addition, the *short-interval capture* case showed a slightly but significantly wider horizontal beam width than the *long-interval capture* case (the t -test, $t(279) = 5.2$, $p < 0.001$), whereas the vertical ones were not significantly different between *short-* and *long-interval capture* cases (the t -test, $t(153) = 1.06$, $p = 0.146$). Figures 3.4A–D also show the directions of the immediate and subsequent prey items relative to the bat's pulse direction accompanying beam patterns while approaching the immediate prey. We found that in the cases of *long-interval captures*, only 30% (59/196 pulses) of pulses covered the subsequent prey within the mean beam width (i.e., $\pm 51 \pm 13^\circ$), whereas the immediate prey was covered by 98% (193/196 pulses) of pulses in the horizontal plane (Fig 3.4A). On the other hand, most of the pulses emitted while approaching the immediate prey in the case of *short-interval captures* (90%, 167/186 pulses) covered the directions of both immediate and subsequent prey items within the mean beam width in the horizontal plane (i.e., $\pm 59 \pm 12^\circ$) (Fig 3.4B). In the vertical plane, the immediate and subsequent prey directions, θ_{ppi} and θ_{pps} , were covered within the vertical mean beam width (i.e., $\pm 27 \pm 14^\circ$) by 94% (164/175 pulses) and 93% (162/175 pulses) of pulses, respectively for *long-interval captures* (Fig 3.4C). In the case of *short-interval captures* (Fig 3.4D), 81% (140/172 pulses) of pulses in the subsequent prey direction, θ_{pcs} , were observed within the vertical mean beam width (i.e., $\pm 25^\circ \pm 8^\circ$). The horizontal mean directions of the subsequent prey ϕ_{pps} in the cases of *long-* and *short-interval captures* were -106° (*long-interval captures*) and 9° (*short-interval captures*), respectively, which were significantly different (the Mardia–Watson–Wheeler test, $B(2) = 165.1$, $p < 0.001$). In addition, those in the vertical plane were also significantly different (the Mardia–Watson–Wheeler test, $B(2) = 32.1$, $p < 0.001$), whereas those were 8° and 7° , respectively. This shows that the bats maintained the subsequent prey within their 3D acoustical field of view in the *short-interval captures*; this supports the acoustical field hypothesis.

Simultaneous coverage of both immediate and subsequent prey items within the

acoustical field of view is supposed to be suitable for planning of the bats' future path to ensure capture of both prey items. To test this, we conducted a numerical simulation (Fig 3.5). The numbers of simulation trials of failure and success were and 4,841 (229,706 pulses) and 1,665 (68,466 pulses), respectively. In the case of failure (only prey 1 was captured), only 39% (89,323/229,706 pulses from 4,841 trials) of pulses covered the subsequent prey within the beam width in the horizontal plane (Fig 3.5A, φ_{pps}). On the other hand, in the success cases, 62% (42,418/68,466 pulses from 1,665 trials) of pulses covered the subsequent prey (prey 2) within the beam width before the capture of the immediate prey (prey 1) in the horizontal plane (Fig 3.5B, φ_{pps}). This also suggests that the bat could potentially succeed in capturing both prey items without invariably keeping the subsequent prey within the beam width in the horizontal plane; that is, the bat could occasionally lose the location of prey 2 just after the start of the simulation trial but could find it while changing the flight direction to capture the prey 1 because the pulse direction was equal to flight direction in the numerical simulation. In the vertical plane, 66% (151,655/229,706 pulses) of pulses in the failure cases and 87% (59,253/68,466 pulses) of pulses in the success cases covered both prey 1 and prey 2 (Figs 3.5C and D). These results suggest that acoustically viewing both the immediate and the subsequent prey items simultaneously increases the success rate for the bats to capture both prey items.

3.4 Discussion

3.4.1 Relationship between acoustic sensing and bat flight

In this study, we found that, in the case of *short-interval capture*, the bats simultaneously maintained both their immediate and subsequent prey items within the horizontal and vertical beam width of the emitted pulse. In addition, the numerical simulations in this study demonstrated that keeping both targets within the beam width increased the success rate of consecutive prey captures. This is considered to be a basic axis of efficient route planning for consecutive-capture flights. These findings suggest that the bats control their pulse direction to cover multiple targets simultaneously within their acoustical field of view. This supports the acoustical field hypothesis derived from our previous studies (Fujioka *et al.*, 2014; Fujioka *et al.*, 2016).

One of our previous studies showed that, in the case of capturing multiple prey

items within a short time interval, the sonar beam of *P. abramus* during foraging shifted directions predictably between the current target and the next target (Fujioka *et al.*, 2014). Based on the measurement results for the 3D beam width in the present study, we found that the bats actually encompassed both prey items in their acoustical field of view. This suggested that when attacking two successive targets, the bats can focus their *acoustic attention* on a subsequent target before capturing the immediate one. This observed acoustic behavior can account for the flight dynamics for the *flight attention* of bats. That is, the bat distributes its *flight attention* between the immediate and subsequent prey so that it can plan its future flight path for a high success rate of consecutive prey captures (Fujioka *et al.*, 2016).

Interestingly, our experimental data show that the bats did not always maintain their immediate prey at the center of the beam width (e.g., Figs 3.3A, B). Instead, they kept the targets within the beam width including the peripheral part, using their wide directional beam effectively to distribute their *acoustic attention* among multiple targets. Such practical operation of wide directional beam scanning employed by the bats is different from the design concepts of existing sensing methods, namely, ultrasonography and radar, which employ high-speed spatial scanning of a narrow directional beam to maintain the spatial resolution of echoes.

Because the movement speed of the prey was insignificant when compared to that of the bats during aerial-feeding flights, at least for a brief period (< 3 s), we set the capture point as the target position during the approach and terminal phases. (See Materials and Methods.) On the other hand, we often see the target prey of *P. abramus* (i.e., mainly small hemipterans and dipterans) (Hirai and Kimura, 2004) swarming at the same position for several tens of seconds, implying that *P. abramus* in our study site may use information on patch locations of insect swarming for consecutive captures. For example, Fig 3.2E shows that the pulse direction corresponded to the positions of Captures 3 and 4, approximately 4 s before Capture 2 occurred. (See from 1.5 s to 3.5 s, in the horizontal axis.) It is still difficult to conclude, however, that the coincidence of the pulse direction with the future prey positions implies that the bat might obtain information on the subsequent prey in advance beyond our expectations. There is another possibility: other prey might be incidentally covered by the acoustic beam before the bat captures its immediate prey during flight. That is, the bat's sonar beam could incidentally cover

multiple prey items while it forages in an area where the prey density is high. Therefore, at the beginning of approaching the consecutive prey, this would be less of a strategy than a coincidence with respect to keeping multiple prey items within the bat's acoustical field of view.

Our previous numerical simulation, however, revealed that a bat's active utilization of positional information on subsequent prey is effective for planning future flight paths during *short-interval captures* (Fujioka *et al.*, 2016). Furthermore, the modeling results in the present study support the idea that maintaining multiple targets within one sonar beam is an effective strategy for consecutive capture. Based on all of these findings, we suggest that the bat's wide sonar beam incidentally covers multiple prey items, and then the bat keeps future targets within its acoustical field of view by actively controlling its acoustical field of view for effective foraging. Mathematical modeling allows us to analyze animal behavior quantitatively, while also suggesting new insights and implementations of cause and effect [e.g., (Vicsek and Zafeiris, 2012; Aihara *et al.*, 2013)]. Further experimental and mathematical investigations into how capable bats are at prediction, and the utilization of the gathered information, will be helpful to clarify the collaborative control of the acoustical field of view and flight path planning employed by bats, in the context of effective adaptation of foraging behavior.

3.4.2 Acoustic sensing in the horizontal and vertical planes

The experimental results demonstrated that the scanning behavior of *P. abramus* differed in the horizontal and vertical planes. We found that the vertical scanning range and the sonar beams were narrower than the horizontal ones in both *long-* and *short-interval capture* flights. (See Figs 3.4A–D.) The bat changed the horizontal acoustical field of view more dynamically than the vertical one, sampling spatial information preferentially in the horizontal plane. In addition, the vertical flight range of the bats (variation of flight height) was also narrower by over five times than the horizontal one. (See Figs 3.2A–C.) Namely, the bats modify their *flight attention* on the horizontal plane for prey that are incidentally on the same vertical plane. Prey insects are usually limited to a certain height above the water. The bats at our study site must avoid bumping into the riverbank during their acrobatic flight maneuvers. Hence, the physical restrictions of the study site may dictate that bats mostly scan their environment horizontally.

On the other hand, the experimental results show that the acoustical coverage rate of subsequent prey in the vertical plane (93% of pulses, Fig 3.4C) was higher than that in the horizontal plane in the case of *long-interval capture* (30%, Fig 3.4A), which is consistent with the numerical simulations of failure cases (capturing only prey 1) (Fig 3.5A vs. Fig 3.5C). Experimental and mathematical results suggest that the bats tend to search for target prey distributed within a certain altitude range, which results to narrow down the vertical scanning and flight ranges. Foraging *P. abramus* often emits pulses downward (see Results) and attacks insect prey in descending flights (Fujioka *et al.*, 2011). Simplifying sensing and flight control in the vertical plane may be effective for complex aerial-feeding flight in 3D space.

3.5 Summary

In this study, we measured the sonar behavior of wild Japanese house bats (*P. abramus*) during natural foraging using large-scale 3D microphone array. The results showed that the bats successively captured multiple airborne insects in short time intervals (less than 1.5 s); they maintained both the immediate and subsequent prey simultaneously within the beam widths of the emitted pulses in both horizontal and vertical planes before capturing the immediate one. Our numerical simulation demonstrated that acoustically viewing both the immediate and subsequent prey simultaneously increases the success rate of capturing both prey items, which is considered to be one of the basic axes of efficient route planning for consecutive capture flight. Our findings suggest that the bats then keep future targets within their acoustical field of view for effective foraging. In addition, in both the experimental results and the numerical simulations, the acoustic sensing and flights of the bats showed narrower vertical than horizontal ranges. This suggests that bats control their acoustic sensing according to different schemes in the horizontal and vertical planes depending on their surroundings. These findings suggest that echolocating bats coordinate their control of the acoustical field of view and flight for consecutive captures in 3D space during natural foraging.

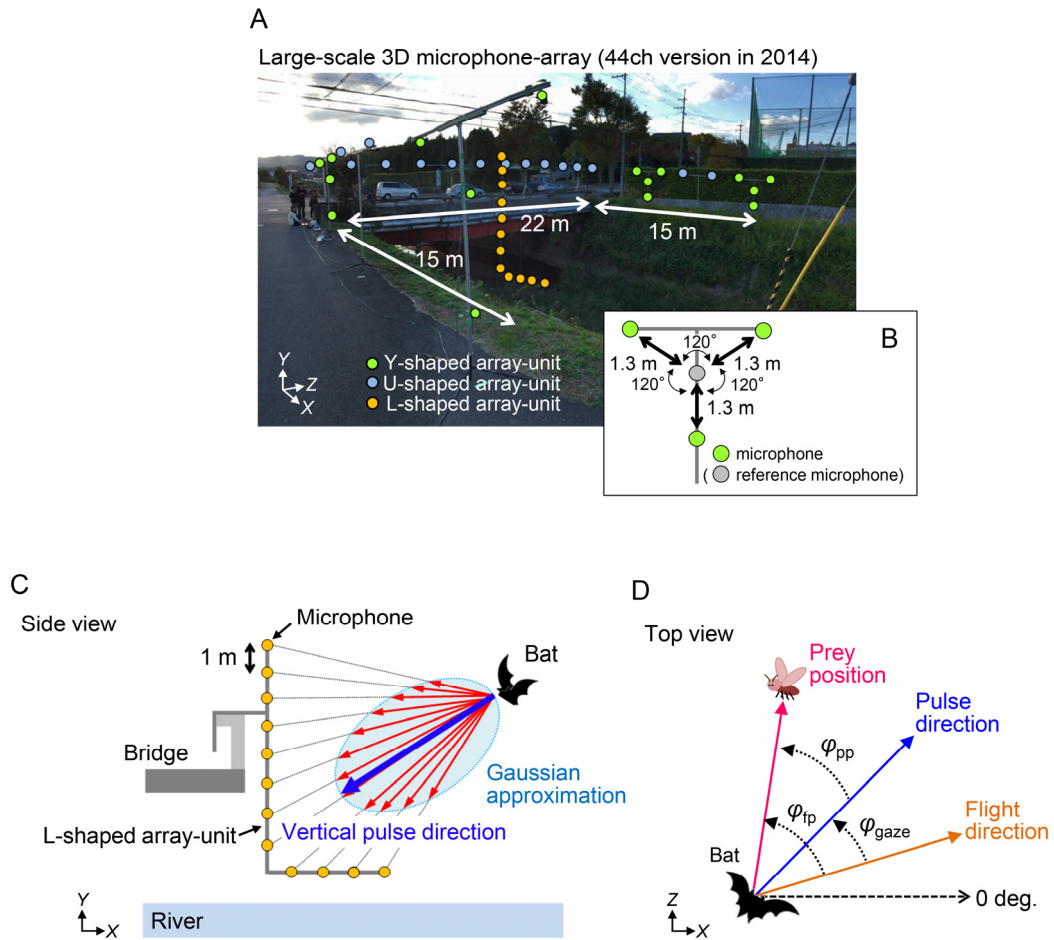


Figure 3.1 Large-scale 3D microphone-array system. (A) Photograph of study site and microphone-array system with 44 microphones consisting of U-shaped 32-ch microphone-array and L-shaped 12-ch microphone-array in 2014. Four Y-shaped arrays (green dots) are part of the U-shaped array. Total of 24 microphones distributed over the entire U-shaped array at the same horizontal level were used to measure horizontal pulse direction, whereas L-shaped array units (orange dots) measured vertical pulse direction. Y-shaped array was used to reconstruct 3D flight paths of the bats. (B) A schematic diagram of the Y-shaped array unit. (C) Side view of the L-shaped array unit. The vertical pulse direction (blue arrow) was determined from the peak of a Gaussian curve (light blue curve), based on the sound pressure vectors (red arrows) across all 12 microphones. The horizontal pulse direction was also determined by the same procedure using the horizontal U-shaped microphone-array. (D) Definitions of the positional relationship between the bat and the target. The gaze angle φ_{gaze} (or θ_{gaze}) was the pulse direction (blue arrow) relative to the flight direction (yellow arrow) of the bat. The directions of the capture positions (prey position) φ_{fp} (or θ_{fp}) and φ_{pp} (or θ_{pp}) were the prey direction (magenta arrow) relative to the flight direction and the pulse direction of the bat, respectively. Here, φ is the horizontal angle and θ is the vertical angle. (See Table 3.1.)

Table 3.1 Definitions of angular variables.

Angular variables	Definitions
φ_{fp}	horizontal prey direction relative to flight direction
θ_{fp}	vertical prey direction relative to flight direction
φ_{pp}	horizontal prey direction relative to pulse direction
θ_{pp}	vertical prey direction relative to pulse direction
φ_{gaze}	horizontal pulse direction relative to flight direction
θ_{gaze}	vertical pulse direction relative to flight direction

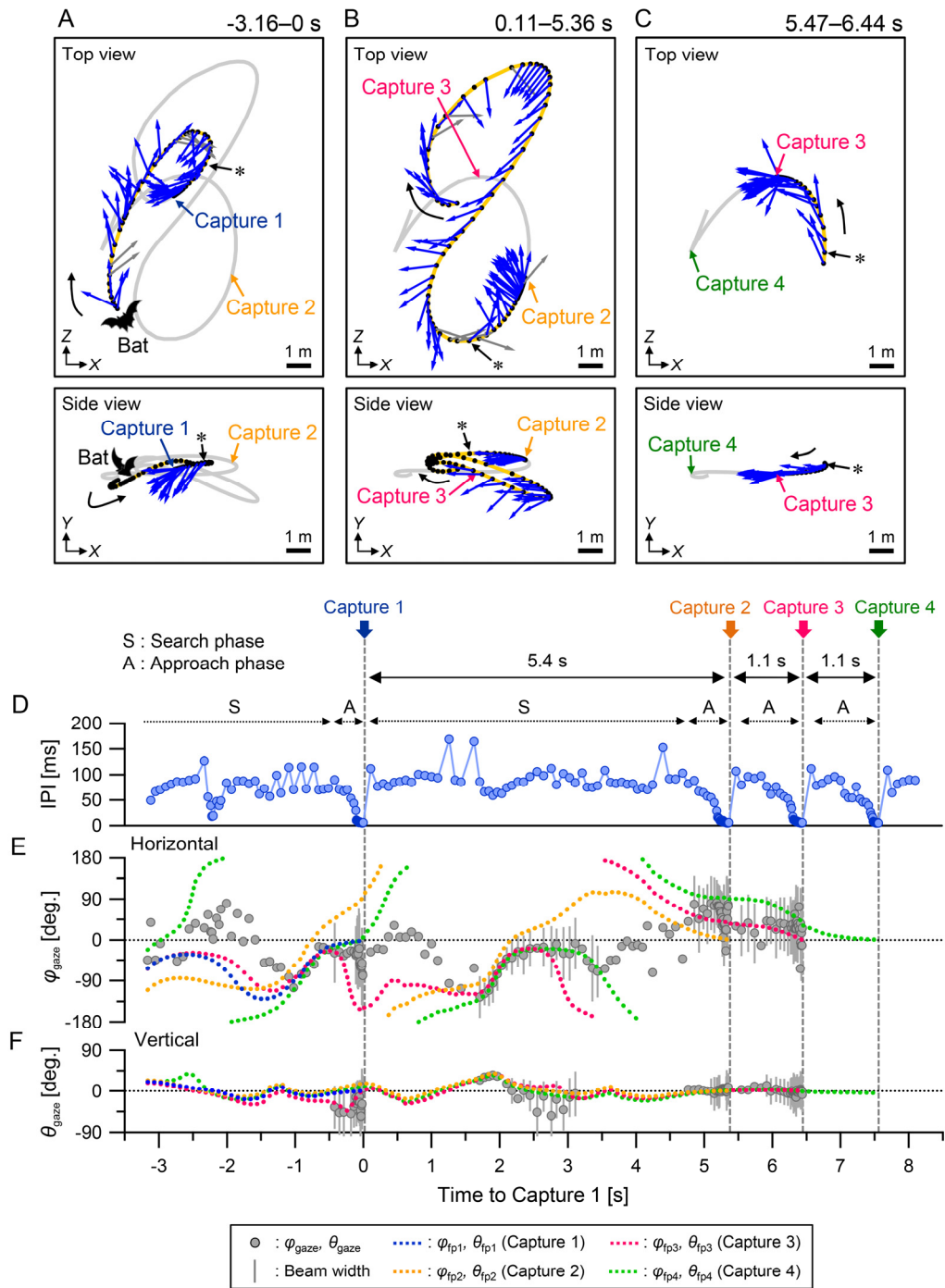


Figure 3.2 Typical example of multiple consecutive capture flight of *P. abramus* in the field. (A–C) Top (top panels) and side (bottom panels) views of the 3D flight path and pulse directions of the bat attacking four successive targets. The observed 3D flight path and pulse directions of the bat were separated into three sections according to the timing of each consecutive target capture; namely, from start to Capture 1 (A), from just after Capture 1 to Capture 2 (B), and from just after Capture 2 to Capture 3 (C). The black curved arrows indicate the initial flight direction of the bat. (to be continued later)

The blue arrows indicate the directions of pulse emission by the bat. The asterisks show the position where the bat started the approach phase. The gray arrows indicate the pulse emitted toward the out of the U-shaped microphone-array in the horizontal plane. Only pulses emitted before the bat captures its immediate prey are shown in the figure. (D–F) Time series data of IPIs (D), gaze angles (φ_{gaze} and θ_{gaze}), and directions of prey positions in the horizontal (E) and vertical planes (F) during this flight. The beam width of the sonar beam is equivalent to the length of gray vertical lines on the gray plots. φ_{fp} (or θ_{fp}) indicates the direction of the prey position relative to the flight direction of the bat (See Table 3.1.).

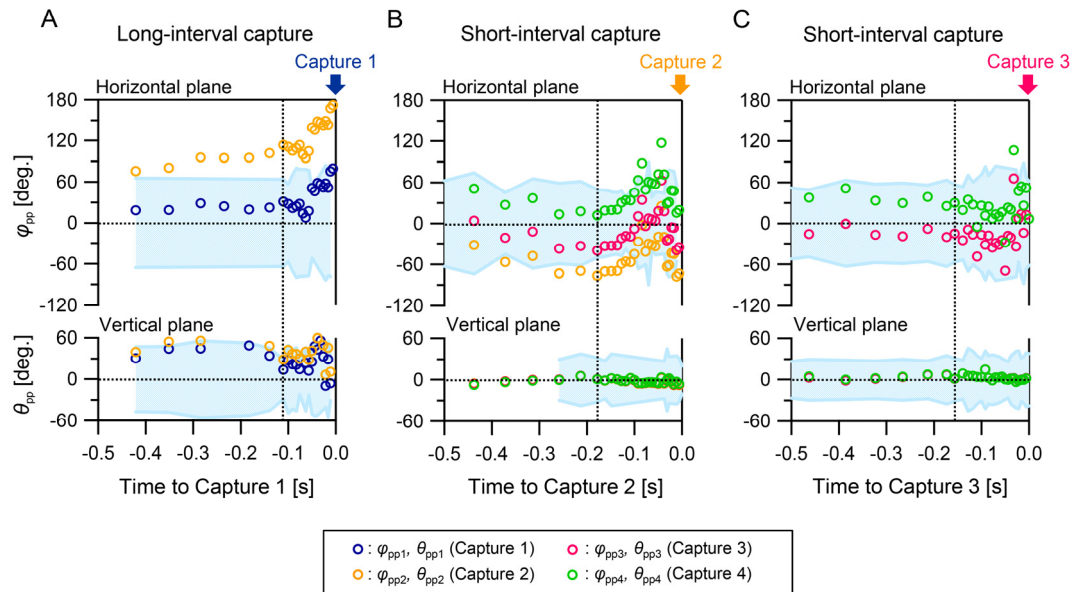


Figure 3.3 Relationship between target positions and bat sonar beam while approaching targets. (A–C) Time series data of φ_{pp} (top panels) and θ_{pp} (bottom panels) during the approach and terminal phases in *long-* (A) and *short-interval captures* (B, C) shown in Figs 2A–C. φ_{pp} (or θ_{pp}) indicates the direction of the prey position relative to the pulse direction. (See Table 3.1.) The light blue areas show the range of the -6 dB beam width of the bat’s sonar beam relative to the bat’s pulse direction. The vertical dashed lines show the timing of transition from the approach phase to the terminal phase.

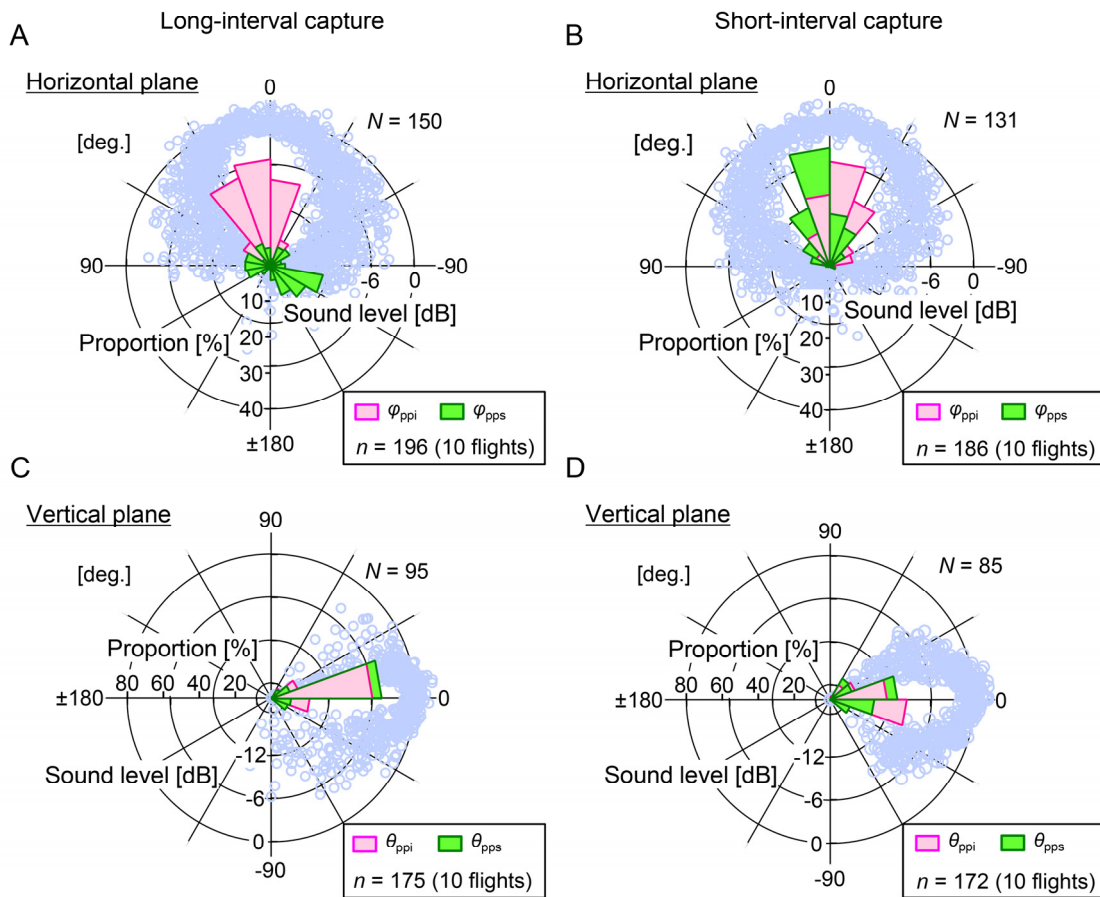


Figure 3.4 Relationship between the prey direction relative to pulse direction and the directivity patterns of the sonar beam as bats converge on immediate prey in *short-* and *long-interval captures*. Circular histograms show the directions of the immediate (horizontal, φ_{ppi} ; vertical, θ_{ppi}) and subsequent (horizontal, φ_{pps} ; vertical, θ_{pps}) prey relative to the bat's pulse direction. Data were obtained from the recording sounds while the bat converged immediate prey (except for buzz II) for 10 flights each of (A, C) *long-* and (B, D) *short-interval captures*. The open light-blue circles indicate the amplitude measured at each microphone, showing the directivity patterns of the sonar beam. We defined 0 dB as the peak value of the curve-fitted directivity pattern of each sound. The vertical (A, B) and horizontal (C, D) axes for the circular histogram show the proportion relative to each number of pulses. Note that the sample sizes in the circular histogram differ from those in the directivity patterns. This is because the beam width data were analyzed only for pulses whose horizontal and vertical beam width could be measured at the same time.

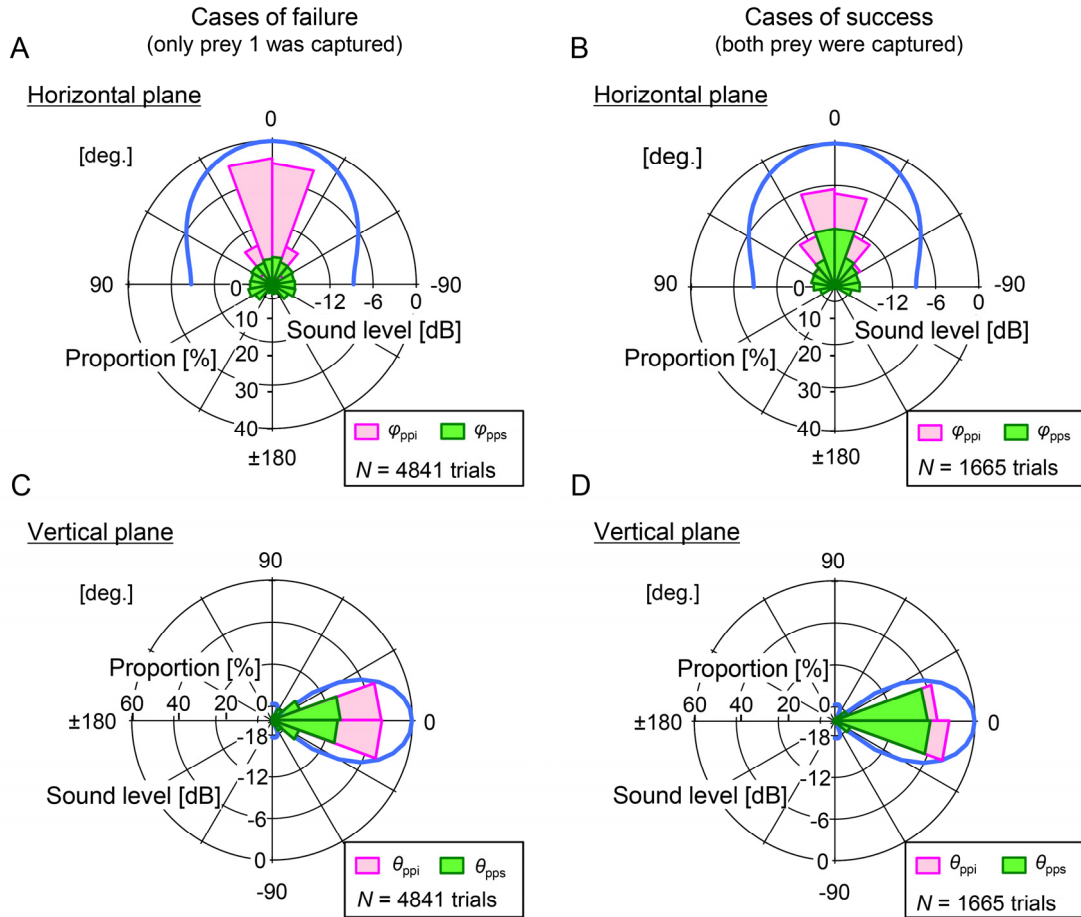


Figure 3.5 Numerical simulation of relationship between prey direction relative to pulse direction and directivity patterns of sonar beam as bats converge on immediate prey during failure and success cases. Circular histograms show the directions of the immediate (horizontal, φ_{ppi} ; vertical, θ_{ppi}) and subsequent (horizontal, φ_{pps} ; vertical, θ_{pps}) prey relative to the bat's pulse direction calculated based on the simulation results. A simulation starts when the bats start to converge on immediate prey, and pulses to obtain the target positions were emitted at every step. A parameter set was defined as a success when the bat captured (close in within 10 cm) the immediate prey (prey 1) and then the subsequent prey (prey 2) in sequence (capturing both prey items), without losing the location of the prey which the bat intend to capture. The simulation was assumed for the phase in which the bat converge immediate prey. Data were taken from 4,841 trials of failure (A, C) and 1,665 trials of success cases (B, D). The light blue lines show the directivity patterns of the sonar beam used in the numerical simulation.

Chapter 4. Mathematical modeling of flight and acoustic dynamics of an echolocating bat during multiple-prey pursuit

(Sumiya *et al.*, 2015)

4.1 Introduction

Nonlinear dynamics of mobile entities (e.g., animals, macromolecules, and robots) attracts a great deal of attention in physics, mathematics, and biology (Vicsek and Zafeiris, 2012). Theoretical and experimental studies on such spatial dynamics are required to reveal the sophisticated mechanisms of motion control of living beings.

Hunting behavior is essential for animals to survive in the wild, because they need to maximize their sensing ability to approach target prey with higher spatial resolution. In particular, bats are unique animals as they detect and capture targets by active sensing using ultrasound. Bats can recognize the physical attributes of their environment with great accuracy by comparing emitted pulses to returning echoes. Using the echolocation strategy, bats can successively capture multiple small moving insects. For example, Japanese house bats (*P. abramus*) fly in large open spaces and capture a few hundred insects per night. It is important for them to sense multiple small insects in order to capture them in a short time interval. Our previous study revealed that bats achieve successive captures in a short time interval of around one second by dynamically changing their flight paths and the acoustic properties of their echolocation sounds (Fujioka *et al.*, 2011; Fujioka *et al.*, 2014; Sumiya *et al.*, 2017), indicating that bats combine flight and acoustic sensing to capture multiple targets. To reveal the sonar strategy of bats, we need to investigate the relationship between flight and acoustic sensing.

The purpose of this study is to examine the nonlinear dynamics inherent in the controls of a bat's flight and pulse directions while approaching successive targets. To achieve this, we propose a mathematical model describing the dynamics of the flight and pulse directions, and then perform numerical simulation to demonstrate that our model can qualitatively explain successive prey captures of an echolocating bat.

4.2 Materials and Methods

4.2.1 Mathematical modeling of flight and acoustic dynamics

In the previous study, Aihara *et al.* (2013) proposed a mathematical model describing the flight dynamics of a bat while approaching a single target in a flight chamber. An extension of the model is required to further examine the mechanism of the pursuit behavior for multiple targets in the field.

In this study, we take into account two points: the bat's flight dynamics and pulse-emission dynamics. Experimental studies using a custom-made microphone array showed that the changes in the flight and pulse directions in the vertical plane are much narrower than in the horizontal plane (Sumiya *et al.*, 2017). Hence, we focus on the flight and pulse-emission dynamics in the horizontal plane, as in our previous study (Aihara *et al.*, 2013). First, the dynamics of the bat's position is modeled as follows (Aihara *et al.*, 2013):

$$\frac{d}{dt} \begin{pmatrix} x_b(t) \\ y_b(t) \end{pmatrix} = v_b \begin{pmatrix} \cos \varphi_f(t) \\ \sin \varphi_f(t) \end{pmatrix}, \quad (1)$$

where $(x_b(t), y_b(t))$ represents the position of the bat in the horizontal plane, and $\varphi_f(t)$ is the flight direction of the bat in the same plane (see Fig. 4.1A). We define $\varphi_f(t)$ as a variable ranging from $-\pi$ to $+\pi$ rad. Parameter v_b is the flight velocity of the bat.

Next, we model the situation that a bat approaches two targets while changing both flight and pulse directions. In this model, $\varphi_p(t)$ represents the horizontal pulse direction, and $\varphi_{bt1}(t)$ and $\varphi_{bt2}(t)$ represent the directions from the bat to Target 1 and Target 2, respectively (see Fig. 4.1A). We assume that the bat can estimate $\varphi_{bt1}(t)$ and $\varphi_{bt2}(t)$ as long as the respective targets are positioned within the beam of the pulse because the real bats localize their targets by using binaural cues from echoes (Aytekin *et al.*, 2004). We defined $\varphi_p(t)$, $\varphi_{bt1}(t)$ and $\varphi_{bt2}(t)$ as variables ranging from $-\pi$ to $+\pi$ rad. The dynamics of the bat's flight direction $\varphi_f(t)$ and pulse direction $\varphi_p(t)$ is modeled as follows:

$$\frac{d\varphi_f(t)}{dt} = \frac{1}{\delta_f} [\alpha_f \sin(\varphi_{bt1}(t) - \varphi_f(t)) + \beta_f \sin(\varphi_{bt2}(t) - \varphi_f(t))], \quad (2)$$

$$\frac{d\varphi_p(t)}{dt} = \frac{1}{\delta_a} [\alpha_a \sin(\varphi_{bt1}(t) - \varphi_p(t)) + \beta_a \sin(\varphi_{bt2}(t) - \varphi_p(t))], \quad (3)$$

where δ_f and δ_a are positive weighting factors, and α_f and β_f (or α_a and β_a) are the parameters minimizing the angular differences between $\varphi_f(t)$ and the target directions (or between $\varphi_p(t)$ and the target directions) (Aihara *et al.*, 2013). Here, α_f and β_f can be understood as the parameters describing flight attention to the targets, while α_a and β_a can be understood as the parameters of acoustic attention to the targets. For example, the model bat shifts $\varphi_f(t)$ (or $\varphi_p(t)$) to Target 1 when α_f (or α_a) takes a positive value (Aihara *et al.*, 2013). Note that the model of Eqs. (2) and (3) can be extended to the model describing the flight and acoustic dynamics for more than two targets by adding another term of sinusoidal function on the right sides.

To study the suitable ratio of flight attention and acoustic attention for successful prey capture, we constrain (α_f, β_f) and (α_a, β_a) as follows:

$$\alpha_f^2 + \beta_f^2 = 1, \quad (4)$$

$$\alpha_a^2 + \beta_a^2 = 1. \quad (5)$$

Consequently, these parameters are described by using two parameters γ_f and γ_a , ranging from $-\pi$ to $+\pi$ as follows:

$$\alpha_f = \sin \gamma_f, \quad (6)$$

$$\beta_f = \cos \gamma_f, \quad (7)$$

$$\alpha_a = \sin \gamma_a, \quad (8)$$

$$\beta_a = \cos \gamma_a. \quad (9)$$

The bat's sonar beam is then modeled as a circular piston oscillating in an infinite baffle based on the previous study (Jakobsen and Surlykke, 2010; Fujioka *et al.*, 2016). The maximum search range (R_{\max} , Fig. 4.1B) was set as 5 m (Fujioka *et al.*, 2011; Fujioka *et al.*, 2016).

4.2.2 Numerical simulation

The dynamics of the horizontal flight direction $\varphi_f(t)$ and pulse direction $\varphi_p(t)$ was numerically calculated based on the present mathematical model of Eqs. (1)–(9).

4.2.2.1 Initial conditions

The initial position and flight direction of the bat was set as $(x_b(t=0), y_b(t=0)) = (0, 0)$ and $\varphi_f(t=0) = 0$. In our dprevious study, the initial positions of prey 1 (Target 1) and prey 2 (Target 2) were randomly determined in every trial in the echolocation distances ranging from 1.2 m to 5.0 m (Fujioka *et al.*, 2016). Further analysis of experimental data showed that the distances from the bat to the immediate target (Target 1) and the subsequent target (Target 2) were around 2 m and 4 m, respectively, at the start point when approaching the immediate target. Therefore, the initial distances from the bat to Target 1 and Target 2 (i.e., $R_{bt1}(t=0)$ and $R_{bt2}(t=0)$ in Fig. 4.1B) were set as 2 m and 4 m, respectively. Furthermore, the initial directions of the two targets relative to the pulse direction (i.e., $\varphi_{pt1}(t=0)$ and $\varphi_{pt2}(t=0)$ in Fig. 4.1B) were randomly determined within a range of $\pm 60\pi/180$ rad that was estimated as possible echolocation range on the basis of the -6 dB beam width of the directivity patterns of the echolocation pulses of Japanese house bats (i.e., *P. abramus*) (Sumiya *et al.*, 2017). The targets such as small insects fly sufficiently slower than the bats (Fujioka *et al.*, 2016; Sumiya *et al.*, 2017), so that two targets are assumed to stay at each initial position in this numerical simulation. Consequently, the horizontal positions of the two targets (x_{t1}, y_{t1}) and (x_{t2}, y_{t2}) are fixed as follows:

$$\begin{pmatrix} x_{t1} \\ y_{t1} \end{pmatrix} = R_{bt1}(t=0) \begin{pmatrix} \cos(\varphi_{pt1}(t=0) + \varphi_p(t=0)) \\ \sin(\varphi_{pt1}(t=0) + \varphi_p(t=0)) \end{pmatrix}, \quad (10)$$

$$\begin{pmatrix} x_{t2} \\ y_{t2} \end{pmatrix} = R_{bt2}(t=0) \begin{pmatrix} \cos(\varphi_{pt2}(t=0) + \varphi_p(t=0)) \\ \sin(\varphi_{pt2}(t=0) + \varphi_p(t=0)) \end{pmatrix}, \quad (11)$$

where the initial pulse direction of the bat (i.e., $\varphi_p(t=0)$) was set to 0.1 rad to avoid to be same as the initial flight direction ($= 0$ rad). The flight velocity of the bat v_b was set as 5 m/s on the basis of experimental results using *P. abramus* (Fujioka *et al.*, 2016).

4.2.2.2 Conditions of prey capture

Successful target capture was defined as a case that the distance from the bat to the targets becomes less than 10 cm without losing the locations of the targets from the bat's echolocation range. Note that the distance 10 cm corresponds to the wing length of *P. abramus* (Hiryu *et al.*, 2007).

4.2.3 Field measurement

Recent studies of biosonar (e.g., in bats and dolphins) have developed in tandem with the microphone-array measurement technology for tracking animal movements based on sound recordings in the field (Fujioka *et al.*, 2011; Seibert *et al.*, 2013; Fujioka *et al.*, 2014; Sumiya *et al.*, 2017). In this section, we show the results of the field measurement of the sonar behaviors of the Japanese house bats *P. abramus* (Vespertilionidae, 10–15 cm wingspan, 5–8 g body mass) (Fujioka *et al.*, 2011; Fujioka *et al.*, 2014; Sumiya *et al.*, 2017). During natural foraging, *P. abramus* emits relatively long pulses (9–11 ms) of shallow swept frequency modulated (FM) signals with the energy concentrated in the terminal sweep frequency of the fundamental component at around 40 kHz (Hiryu *et al.*, 2008a). The echolocation sounds of *P. abramus* were recorded by using a large-scale 3D microphone array with 500 kHz sampling rate (Fig. 4.2A) at a stream near the campus of Doshisha University in southern Kyoto Prefecture, Japan, from the early summer to the fall during the evenings (Sumiya *et al.*, 2017). The 3D locations of the bats were obtained using time difference of arrival (TDOA) between a reference and three other microphones (which were separated by 1.3 ± 0.01 m in the Y-shaped array, as shown in Fig. 4.2B) (Au and Herzing, 2003). The locations where the bats captured the targets were determined based on the occurrence of feeding buzzes (i.e., the successive pulse emissions with extremely short time intervals just before capturing the focal target) (Kalko and Schnitzler, 1989; Hiryu *et al.*, 2008a).

4.3 Results

4.3.1 Results of simulation

Figures 4.3A and B show the results of our numerical simulation representing the successful examples of two consecutive captures by a model bat. The gray line describes the flight path of the model bat. The black thin arrows represent the pulse directions $\varphi_p(t)$. All four variables in this model (i.e., $x_b(t)$, $y_b(t)$, $\varphi_f(t)$, and $\varphi_p(t)$) were calculated by using the fourth-order Runge–Kutta method with a time step of 0.001 s.

Numerical simulation shows that the model bat successfully captured both of the targets, without losing them from its echolocation range, when the model bat directed $\varphi_p(t)$ to Target 2 rather than Target 1 (i.e., $\gamma_a = 0.2\pi$ or -0.1π ; see Figs. 3A and B). In contrast, the model bat could not capture the two targets when $\varphi_p(t)$ was directed only to Target 1 (i.e., $\gamma_a = 0.5\pi$; see Fig. 4.3C).

These results at specific parameter values suggest that, for successive prey captures in a short time interval, it is important to emit pulses toward the subsequent target before capturing the immediate target.

4.3.2 Flight behavior of the bat while approaching multiple targets

Figure 4.4 shows an example of the top view (i.e., horizontal plane) of the 3D flight path and pulse directions of a bat during natural hunting measured by the large-scale 3D microphone array (Figs. 4.2A and B). In this case, the bat successively captured two insects (Targets 1 and 2) in a short time interval of 1.1 s, directing its pulses toward Target 2 before capturing Target 1 (Fig. 4.4), which is consistent with the results of our previous experimental study (Fujioka *et al.*, 2014) and also consistent with the results of the present numerical simulation (Figs. 4.3A and B).

4.4 Discussion

Our numerical simulation demonstrated that the present model can qualitatively explain successive prey capture with specific parameter values. In addition, it is suggested that such successive prey capture is accomplished when a model bat flies toward the immediate target while emitting pulses toward the subsequent target.

Comparison between the theoretical and experimental studies will allow us to

examine the detailed mechanisms of multiple-prey capture according to the present mathematical model. For this purpose, the parameters φ_f and φ_a , which represent flight attention and acoustic attention, respectively, need to be estimated from the experimental data using the microphone-array system. We then need to conduct the numerical simulation by changing the parameter values and initial conditions, and also need to compare the experimental and theoretical results to clarify the detailed dynamics of the decision-making of the bats during natural hunting.

It remains as a future problem to compare the dynamics of the present model to that of related mathematical models. Previous modeling studies of bat's pursuit behavior suggests that bats use a functionally predictive flight strategy during chasing erratically moving insects in the flight chamber (Ghose *et al.*, 2006), whereas another modeling study suggests that bats can successfully capture insects using nonpredictive strategy (Kuc, 1994). In addition, the relationship between the visual line and motion control of living beings such as humans or insects is associated with the bat's sonar behavior from the viewpoint of the interaction between attention and motion control (Land and Tatler, 2001; Ghose and Moss, 2006). Comparison between our model and these related models would be helpful to investigate the validity and generality of our model.

In this study, we propose a mathematical model describing flight and pulse-emission dynamics in the horizontal plane. However, real bats fly around in the 3D space, sensing their surrounding objects. An extension of the present model will be required in order to study flight and echolocation mechanisms in the 3D space, such as the mechanism using descending motion during prey pursuit. In fact, we have already extended the model of bat's flight direction in horizontal plane to the model in the 3D space (Fujioka *et al.*, 2016). The pulse-emission dynamics would be extended to the 3D space by the similar approach. The extension of bat's sonar-beam shape to the 3D space is also required according to the previous studies (Kuc, 1994; Jakobsen and Surlykke, 2010; Kuc, 2012).

The real bats detect the echoes from the targets with a certain time delay varying depending on the distance between the bat and the target (Liang and Palakal, 1997). Such a time delay would play an important role both for echolocation and motion control. The model of Eqs. (2) and (3) can be modified to take such a time-delay effect into account by replacing $\varphi_{bt1}(t)$ and $\varphi_{bt2}(t)$ with $\varphi_{bt1}(t-\tau)$ and $\varphi_{bt2}(t-\tau)$, using a variable τ representing

the time delay.

4.5 Summary

We proposed a new mathematical model describing the nonlinear dynamics of the flight and pulse directions of an echolocating bat approaching two successive targets. Numerical simulation of the present model shows that the model bat successfully captures both targets within a short time interval without losing them from its sonar beam at specific parameter values. The simulation also suggests that the successive prey capture is completed when the echolocation pulses are directed to the subsequent target before capturing the immediate target. Such a relationship between the flight and acoustic sensing can be also observed in the behavioral data of wild bats.

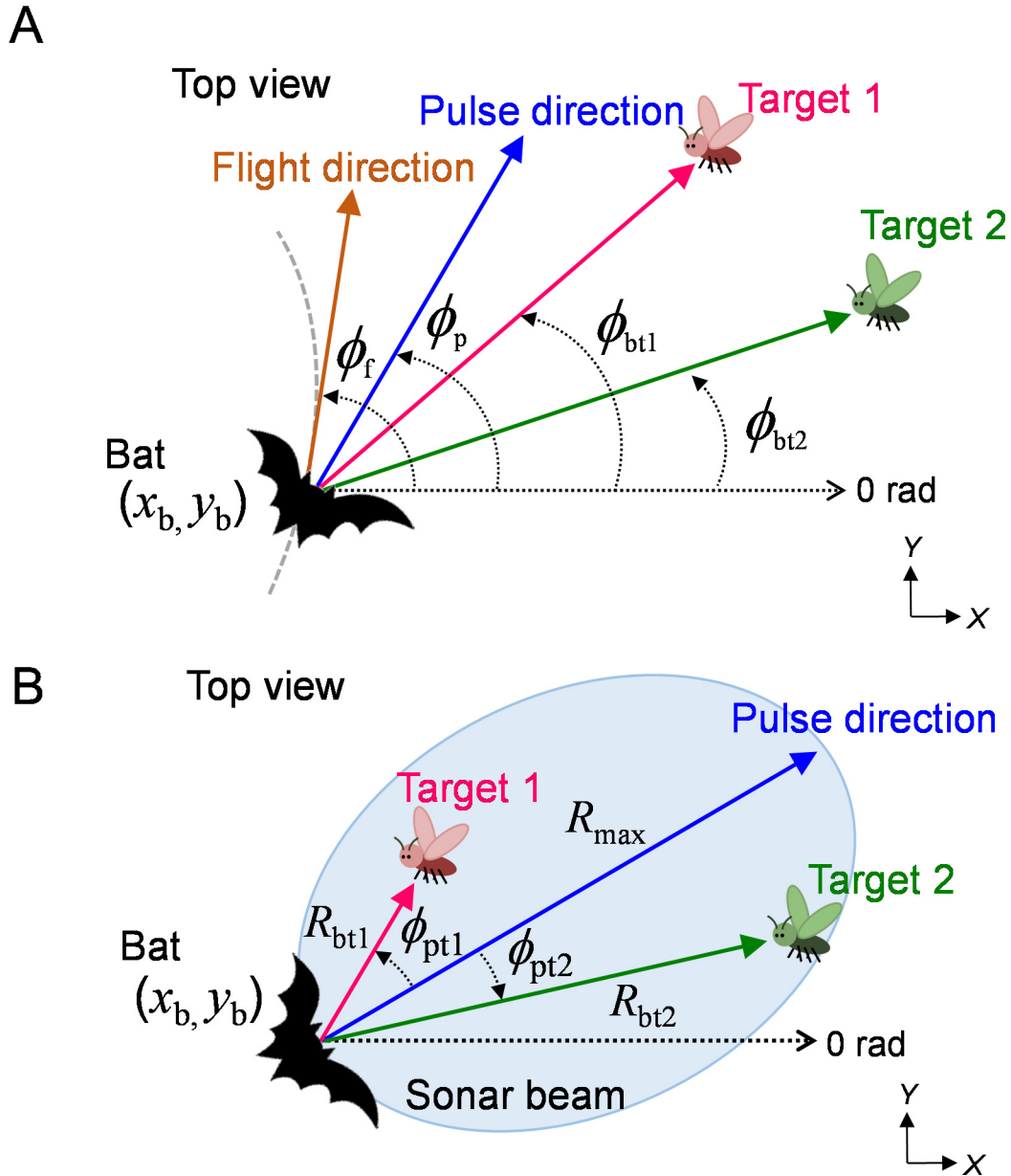


Figure 4.1 Schematic diagram of the present mathematical model described by Eqs. (1)–(3). (A) The relationships between the bat and two targets in the horizontal plane. The model bat controls its flight and pulse directions ($\phi_f(t)$ and $\phi_p(t)$) by sensing the directions from itself to the targets ($\phi_{bt1}(t)$ and $\phi_{bt2}(t)$). (B) Schematic diagram of the sonar beam of the bat. The sonar beam is modeled as a circular piston oscillating in an infinite baffle with a range of 5 m (i.e., $R_{max} = 5$ m).

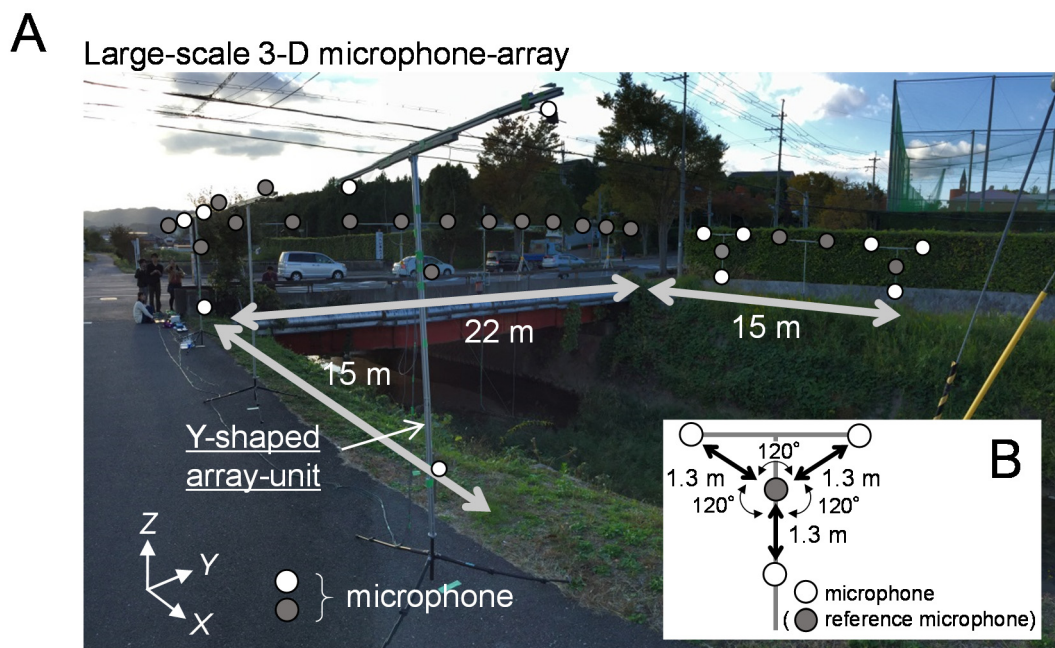


Figure 4.2 Microphone array system in the field. (A) Photograph of the microphone array system in the field. The white circles show the microphones installed at the Y-shaped array units used to calculate the 3D sound coordinates of the bat. (B) A schematic diagram of the Y-shaped microphone array. The distance between the reference microphone (dark gray circle) and each of the three microphones (white circles) was 1.3 ± 0.01 m, and they were distributed with an angular separation of 120° . The 3D flight paths of the bats were reconstructed from differences in the arrival times of the ultrasound pulses among the microphones.

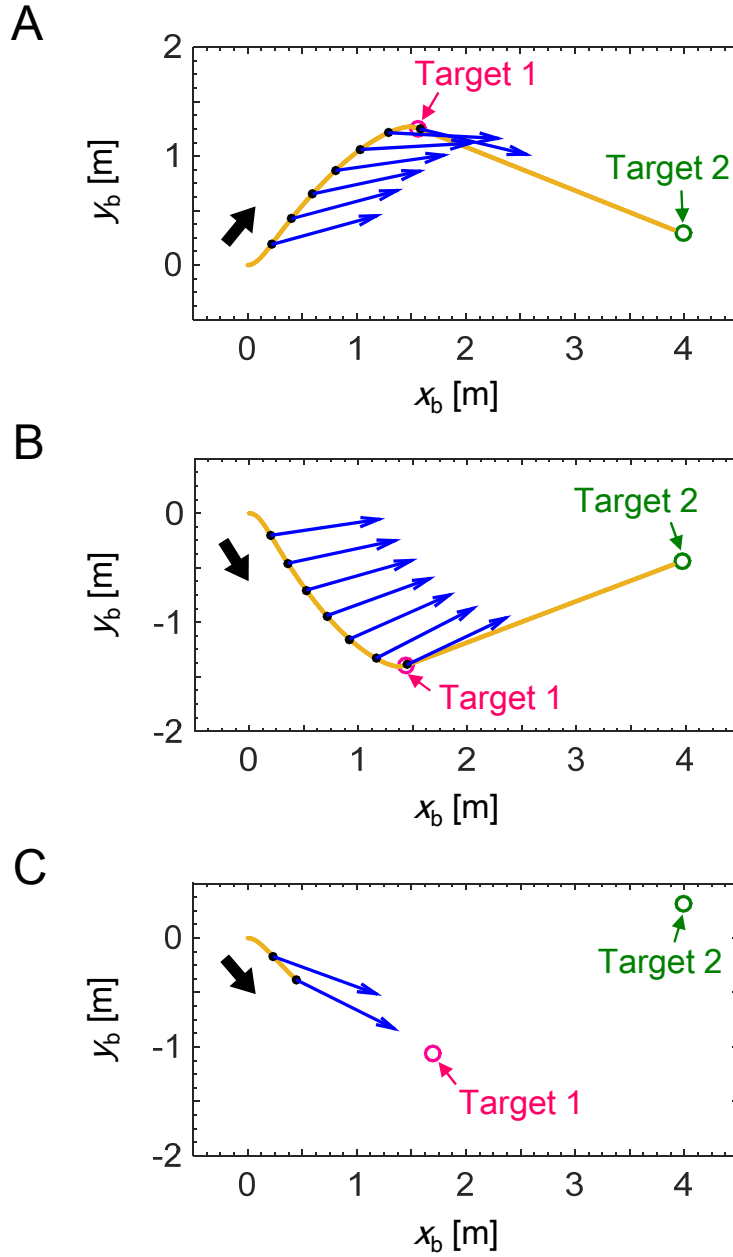


Figure 4.3 Representative cases of the flight path and pulse direction calculated from the present model of Eqs. (1)–(9). The parameter values of the model were fixed as $(\gamma_f, \gamma_a) = (0.6\pi, 0.2\pi)$ (A), $(\gamma_f, \gamma_a) = (0.6\pi, -0.1\pi)$ (B) or $(\gamma_f, \gamma_a) = (0.6\pi, 0.5\pi)$ (C), with $(\delta_f, \delta_a) = (0.01, 0.05)$ and $v_b = 5$ m/s. Gray lines show the flight path of the bat. Black filled arrows show the flight direction. Open circles show the positions of Target 1 and Target 2. The pulse direction $\varphi_p(t)$ showed with black thin arrows is plotted at the intervals of 0.06 s in this figure, to clearly show $\varphi_p(t)$ without overlap. In (A) and (B), the model bat successfully captured both the two targets. In (C), the model bat lost Target 2 from the sonar beam on the approach to Target 1. We judged such a case of losing a target from the echolocation range as a failure to capture in our simulation.

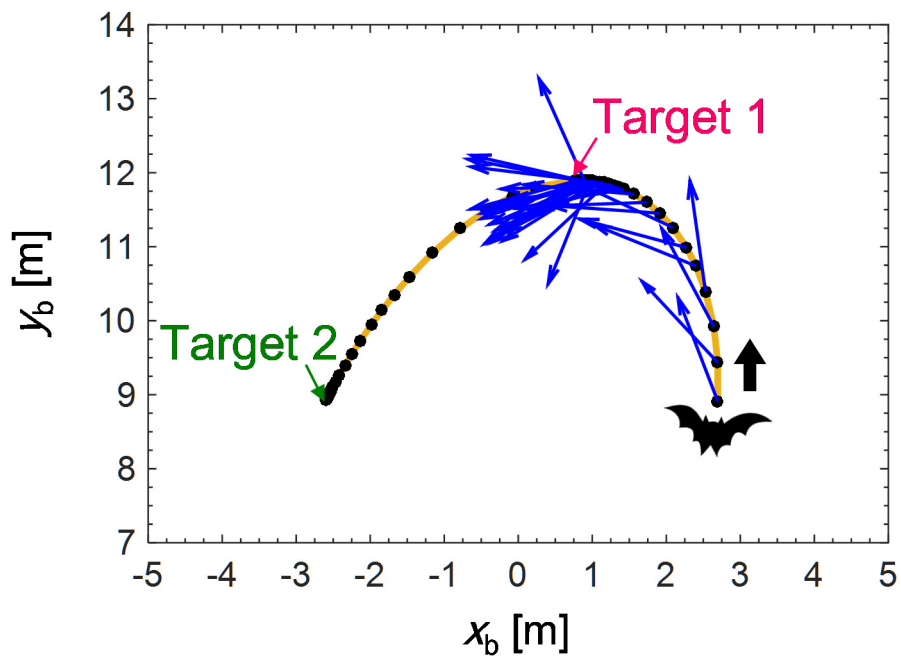


Figure 4.4 Representative case of the flight path (gray line) and pulse directions (black thin arrows) during successive captures of two targets. Black dots show the position where the bat emitted the pulse. Black filled arrows show the flight direction. The time interval to capture two successive targets was 1.1 s.

Chapter 5. Human echolocation using ultrasonic binaural echoes

5.1 Introduction

Bats and dolphins recognize the distance, direction, and shape of an object by echolocation. For example, bats have the ability to detect time delay with high resolution (less than 1 μ s) (Simmons, 1979). In addition, bats can discriminate the category of prey (von der Emde and Schnitzler, 1990) and the texture of an object (Falk *et al.*, 2011) based on the acoustic information of echoes. Moreover, it is also known that dolphins can discriminate the differences between metallic materials and slight differences in thickness by echo location (Hammer Jr and Au, 1980). Dolphins, using their auditory sense and vision, can select a novel object perceived by echolocation under visual conditions; thus, it is thought that dolphins integrate auditory information acquired by echolocation and visual information in the brain (Harley *et al.*, 2003). This suggests that the ultrasonic sensing performed by a biosonar may have the ability to “see by sound,” with a resolution equal to or higher than that of vision. However, the neurophysiological base supporting this advanced cognition has not yet been elucidated. Furthermore, because animals require complicated psychological experiments, the relationship between the acoustic properties of echoes, which bats and dolphins use as acoustic cues, and the perceived information has not yet been clarified.

It has long been reported that some blind people can also sense the surrounding environment by emitting mouth clicks and listening to their echoes (Rice, 1967; Hausfeld *et al.*, 1982). With respect to engineering applications, research on assistive devices for visually impaired people has been performed (Kay, 1974; Ifukube *et al.*, 1991). In addition, participation of a few biological sonar researchers have been conducted (DeLong *et al.*, 2007a; DeLong *et al.*, 2007b). In recent years, brain function measurements by fMRI have indicated that blind people performing echolocation have cerebral visual and auditory cortex activation (Thaler *et al.*, 2011). As a result, attention is being paid by neuroscience to the elucidation of the human echolocation ability.

Based on this background, we approach this study, “what can be seen from the echo?” by focusing on the echolocation ability of humans, with whom it is possible to

conduct psychoacoustic experiments and brain functional measurements. In our psychoacoustic experiments, we design acoustic parameters based on our knowledge of bats, obtained in previous studies (Fujioka *et al.*, 2011). Our ultimate goal is to clarify the biological sonar mechanism and to acquire, by learning from biological mechanisms, new engineering knowledge that will be useful for the sensing technology. Because ultrasound has higher spatial resolution, it is considered as advantageous for echolocation. Therefore, we established a psychoacoustic experimental system using ultrasonic binaural echoes measured by MDH which can reproduce the acoustic space information measured in the ultrasonic range in three dimensions. This will allow humans to perform echolocation by using ultrasound with high spatial resolution, as that performed by bats and dolphins. By pitch-converting the echo binaurally recorded in the ultrasonic range using the $1/n$ scaled MDH to the $1/n$ audible range, an echo including spatial information sensed by ultrasound can be presented to humans. In our previous study, we demonstrated that the sounds captured by the proposed MDH system were more accurately localized outside the head, by the participants, than the normal stereo sounds (Uchibori *et al.*, 2015).

In this study, we measured the echolocation ability using ultrasonic binaural echoes measured by MDH in sighted echolocation novices. In addition, we examined the relationship between echolocation ability and the acoustic characteristics included by the echo. In an echolocation study targeting blind individuals, objects existing in daily life were often used as sensing targets (Milne *et al.*, 2015b), suggesting that blind people should discriminate such targets for their daily lives. In this experiment, as the first echolocation study using ultrasonic binaural echoes measured by MDH, we examined the echolocation ability for targets having various materials and surface structures that are present in daily life. Namely, we tried to measure the ultrasonic binaural echoes from these targets. Then, we conducted a discrimination experiment using the sound stimuli that were created by pitch-converting the ultrasonic binaural echoes for the sighted echolocation novices. In addition, we examined the utility of ultrasounds by examining the relationship between the psychoacoustic experiment results and the acoustic analysis.

5.2 Materials and Methods

5.2.1 Participants

A total of eight participants, aged between 23 and 31 years (Mean \pm SD; 24.9 \pm 2.6), took part in the target discrimination experiment at Doshisha University. All participants reported having normal hearing and no history of hearing difficulties. All participants reported not to have prior experience with echolocation, and they reported to have either normal or corrected normal vision. The testing took place in a sound-attenuated chamber (2.3 m (H) \times 1.6 m (L) \times 1.4 m (W)). The testing procedures were approved by the University ethics board and the participants provided informed consent.

5.2.2 Targets

Figure 5.1a–f shows six custom made targets (a: acrylic board, b: wood, c: foamed polystyrene, d: artificial grass, e: artificial foliage, f: gardening fence), which were used as sound-reflecting objects in the target-discrimination experiment. In addition, we measured the echo without target.

5.2.3 Miniature dummy head

The 1/7 scaled MDH with torso was made from acrylonitrile-butadiene-styrene (ABS) resin using a 3D printer (UP!3D Printer; Techno Solutions, Tokyo, Japan). We used the 3D shape data of a standard dummy head (4128C Head and Torso Simulator, HATS; Brüel & Kjær, Nærum, Denmark), scanned using a 3D scanner (Model 2020i Desktop Laser Scanner; NextEngine, CA, USA) (Uchibori *et al.*, 2015). Because the HATS torso was too large for the 3D scanner, the shape of the torso was not scanned, but it was made so that it roughly imitated the shape of the HATS torso. The width, height, and depth of the MDH torso were 45 mm, 58 mm, and 20 mm, respectively. Omnidirectional condenser microphones of 2.7 mm diameter (B6 Omnidirectional Lavalier; Countryman Associates, Inc., California, USA) were inserted at the entrance of the left and right ear canals of the MDH. The interaural distance of the MDH was 22 mm. In order to present the binaural sounds captured by the MDH to the listeners, the frequency of the binaural sounds captured by the MDH was shifted, with the relative reduction of the head size, by the time-expansion method using an audio editing software

(Audition CC 2015; Adobe). This allows the listeners to hear the pitch-converted binaural sounds, with the acoustic information included in the high frequency range, as audible sounds. In addition, the listeners can perceive 3D sounds when they listen the pitch-converted binaural sounds through headphones (Uchibori *et al.*, 2015).

5.2.4 Sound stimuli

Sound stimuli were the pitch-converted ultrasonic binaural echoes measured by the 1/7 scaled MDH. We measured the ultrasonic binaural echo from each target in an anechoic chamber (3.6 m (W) × 4.4 m (L) × 3.4 m (H)) at Doshisha University by using the experimental system shown in Fig. 5.2. The target was positioned at a distance of 40 cm from the loudspeaker (LS-301-CH; TEAC CORPORATION, Tokyo, Japan). The bats that employ FM sounds use downward FM sounds, while those using CF-FM sounds mainly use a combination of upward FM, CF, and downward FM sounds. Therefore, in this study, we used down FM sweep (7–35 kHz), up FM sweep (7–35 kHz), and CF bursts (7 kHz) as echolocation sound. The direct sound from the loudspeaker and the echoes from the target were measured by the two microphones in the 1/7 scaled MDH. Bats are roughly divided into species using FM sounds and species using CF-FM sounds. In addition, we measured in the same way the direct sound and echoes without the target. The captured signals were digitalized with 32-bit accuracy at a sampling rate of 192 kHz by using a high-speed USB audio interface (OCTA-CAPTURE; Roland, Shizuoka, Japan). The digitalized signals were synchronously stored by using a personal computer (VAIO Z; VAIO Corporation, Nagano, Japan) and by using the audio editing software⁵⁵. Note that the frequency characteristics of loudspeakers and microphones are corrected so that they are within ± 3 dB in the measurement band.

5.2.5 Procedure

Two different randomly selected sound stimuli were presented in a two-alternative forced-choice task (2AFC) by using a custom-made program, EXPLAB (free software for computational experiment), through headphones (MDR-CD900ST, Sony, Tokyo, Japan). The time interval between the sound stimuli (i.e., inter-stimulus interval) was 300 ms. The peak level of amplitude of the direct sound from the loudspeaker was standardized. In the experiment, the participants were required to accurately identify

whether or not the two kinds of signals are the same by writing the answer on the paper.

We conducted four two-test sessions (without feedback of the answer on the display) per participant. There was a total of 168 trials (28 target pairs (number of combination with repetition for choosing 2 targets from all 7 targets) \times 3 signals \times 2 repetitions) in the test session. After completing the experiment, a questionnaire survey on the acoustic cues was conducted for each participant.

5.2.6 Statistical analysis

Non-metric multidimensional scaling (NMDS) was used to evaluate how seven echoes from each object were perceived as different from each other. We calculated the coordinate of each echo on the orthogonal axis in the two-dimensional space consisting of dimension 1 and dimension 2. The average percentage of correct answers shown in Table 1 was taken as the mutual psychological dissimilarity between echoes. All analyses were performed using MATLAB.

A generalized linear mixed model (GLMM) (Schall, 1991) with a binomial distribution and logit link function was used to evaluate the effect of the degree of similarities in the acoustic parameters (i.e., amplitude wave pattern, frequency spectrum, and spectrogram) on the percentage of correct answers in each target pair. When multicollinearity was confirmed in the relationship between the acoustic parameters (the Pearson's correlation coefficient exceeded 0.5), we excluded one of the acoustic parameters from the analysis. The participant's ID was treated as an explanatory variable describing a random effect. The number of correct answers was treated as the response variable. All analyses were performed using R Statistical Software Version 3.2.3. The analyses with GLMMs were conducted using the system function "glmmML".

5.3 Results

Table 5.1 shows the average percentage of correct answers of all eight participants in each target pair. As listed in Table 5.1, the average percentage of correct answers in the case of extremely-different surface conditions (e.g., acrylic board vs. artificial foliage) was more than 90 % (more than the chance level, 50 %), while in the slightly-different surface conditions (e.g., acrylic board vs. foamed polystyrene) it was

under 40 %. Furthermore, the average percentage of correct answers under the downward and upward FM sweep conditions were above chance level (downward FM, 75.4 ± 22.0 ; upward FM, 78.1 ± 24.1 ; CF, 61.8 ± 24.1). The GLMM evaluation showed that the downward and upward FM sweep conditions had a significant positive effect on the average percent of correct answers (downward FM, $\beta = 0.9427 \pm 0.1234$, $z = 7.639$, $p < 0.001$; upward FM, $\beta = 0.9775 \pm 0.1240$, $z = 7.882$, $p < 0.001$). Additionally, the average percentage of correct answers 5–10 % was higher in the second trial than in the first trial.

The distances between the stimuli calculated based on NMDS, shown in Figs. 5.4A–C, correspond to perceptual distances (i.e., how much similar they sound). Although the distances between the plots of acrylic, wood, and polystyrene foam (a–c) were relatively close under all signal conditions, the plots were more dispersed under downward and upward FM sweep conditions than under CF burst condition. Under CF burst condition, the plots were divided into two groups (a–d and e–g), and the distances of plots among targets were approximately zero in both groups.

Figures 5.5A–C show the frequency spectra of the echoes of acrylic board (a), wood (b), and foamed polystyrene (c), which tended to have relatively high psychological similarity as shown in Fig. 5.4. As depicted in Fig. 5.5A–C, the amplitude spectra of the echoes of acrylic board (a) and wood (b) exhibited almost the same spectral pattern under all signal conditions. However, the amplitude spectrum of wood (b) and foamed polystyrene (c) showed a slightly different notch pattern under downward and upward FM sweep conditions, whereas such slight differences were not observed under the CF burst condition.

5.4 Discussion

5.4.1 Discrimination performance using ultrasonic binaural echoes

In the pair of wood (b) and foamed polystyrene (c) (Figs. 5.5A–C), the difference in the notch pattern seen in the amplitude spectrum was clear, and the difference between the correct answer rates of FM sound and CF sound was remarkable. Even in the interview survey of the subjects, most of participants judged based on a slight difference in timbre. This suggests that even if it is difficult to discriminate by CF sound, it can be easily supposed that FM sound is effective in discrimination because it can use a spectrum

pattern as acoustic cue. Therefore, we calculated the coordinates of each echo on the direct axis in the two-dimensional space. They were calculated using the same method as in Fig. 5.4, using the peak value of the cross correlation of the amplitude spectrum of each echo as the acoustic similarity. The results are shown in Figs. 5.6A–C. As shown in Figures 5.6A–C, the acoustic similarity is high in a combination with low correct answer rate (e.g., wood (b) and foamed polystyrene (c)). On the contrary, it can be seen that the acoustic similarity is low in combination with a high correct answer rate (e.g., wood (b) and artificial foliage (e)).

Table 5.2 shows the result of the GLMM for evaluating the effect of target discrimination performances and acoustic similarity of the echoes ($N = 168$). The differences in acoustic parameters among the targets had a significant negative effect on the number of correct answers under each signal condition (Table 5.2). This indicates that the echo discrimination rate decreases when the acoustic similarity is high. In the future, we will continue to examine the signal parameters that are useful for discriminating the physical features of the target based on the acoustic similarity of the echo.

5.4.2 Comparison with the echolocation performance using mouth clicks

Finally, we carried out an echolocation experiment using mouth clicks emitted by the four participants (e.g., Fig. 5.7) as in previous human echolocation studies (Kellog, 1962; Rice, 1967). In this additional experiment, the participants were required to identify whether two targets randomly presented were the same or not, based on a 2AFC task. The testing was performed in an anechoic chamber (3.6 m (W) \times 4.4 m (L) \times 3.4 m (H)) at Doshisha University. The target was same as that used in the discrimination experiment employing ultrasonic binaural echoes. The targets were positioned in isolation directly ahead of the participant, with the center at ear level. Once the trial began, the participants were given five opportunities to scan the object using their mouth clicks. There was a total of 28 trials (28 target pairs \times 1 repetitions) in the test session. We conducted four test sessions per participant over two days. Figure 5.8 shows the average percentage of correct answers of the four participants under ultrasonic binaural echo conditions (downward and upward FM sweep, and CF burst conditions) and mouth click condition. The discrimination performance using mouth clicks was approximately 11 % lower (56.9 ± 27.3 %) than that using CF ultrasonic binaural echoes (68.1 ± 30.8 %). This is because

the peak frequency wavelength of the mouth click was shorter than that of the ultrasonic binaural echoes. In addition, the ultrasonic binaural echoes measured using MDH are presented to the participants after pitch conversion, which means that the time axis was extended according to the MDH scale. However, during natural human echolocation using mouth clicks, the participants listen to the echoes from the target in real time. Because the signal length of the mouth click is approximately 5 ms (Fig. 5.7), it is conceivable that the echo from the target ahead of 40 cm is overlapped with the mouth click emitted by themselves, and thus, it is considered that the difficulty level of echo discrimination increases. This suggests that the ultrasonic binaural echo with high frequency and temporal resolution may be more advantageous than the mouth click for discrimination of surface structure and material of a nearby object.

However, the sound pressure might not be sufficient to hear the echo because the participants had no experience in human echolocation using mouth clicks. Thaler and Castillo-Serrano (2016) demonstrated that even sighted participants could perform well as blind echolocation experts when they used a loudspeaker to emit the mouth clicks, instead of emitting these by themselves (Thaler and Castillo-Serrano, 2016). This suggests that the participants might perform better by using artificial mouth clicks. Therefore, in order to examine the relationship between frequency of the signal and echolocation performance, it is necessary to conduct, in a future work, a psychological experiment using the same artificial sound.

5.5 Summary

We showed research results on the establishment of a human echolocation system that takes advantage of bat echolocation. Based on the proposed human echolocation system, we investigated the discrimination ability of sighted subjects for object texture to understand acoustic cues for texture recognition in human echolocation. Our results suggested that it is possible to discriminate targets with extremely different surface structures and materials, which are easy to discriminate by vision, by listening to ultrasonic binaural echoes. Furthermore, the rate of correct answers in the CF sound condition was approximately 13% lower than those in the FM sound condition. The correlation diagram among targets by multidimensional scaling was dispersed more

remarkably in the FM sound condition. When the target pair had slightly different surface conditions, differences in the notch pattern of amplitude spectra were observed especially in the FM sound condition. These suggest that FM ultrasonic binaural sound is more effective for slightly different texture perception than CF ultrasonic binaural sound.

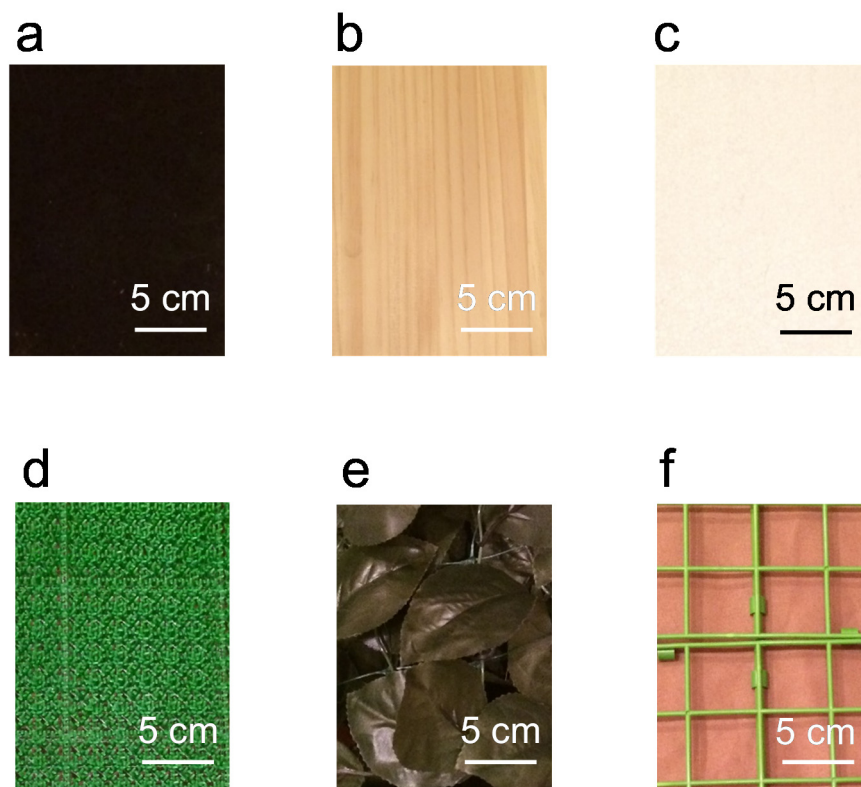


Figure 5.1 Photographs of six targets (a: acrylic board, b: wood, c: foamed polystyrene, d: artificial grass, e: artificial foliage, f: gardening fence).

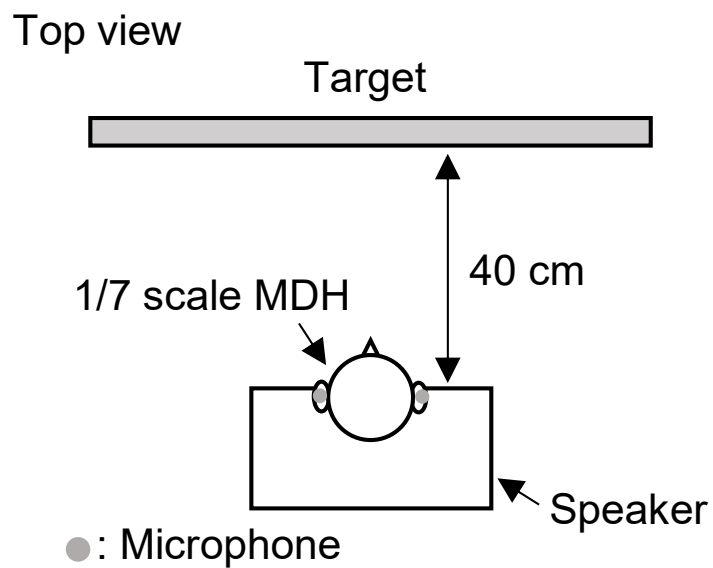


Figure 5.2 Schematic diagram of the experimental system. The distance between the 1/7 scaled MDH attached to the loudspeaker and the target was 40 cm.

Table 5.1 Average percent of correct answers of the eight participants in each target pair under the downward FM (top), upward FM (middle), CF (bottom) signal conditions (a: acrylic board, b: wood, c: foamed polystyrene, d: artificial grass, e: artificial foliage, f: gardening fence, g: anechoic chamber).

Down FM		<i>N</i> = 8 participants					
	a	b	c	d	e	f	g
a	81.3	40.6	37.5	68.8	100.0	96.9	100.0
b		62.5	40.6	84.4	96.9	100.0	100.0
c			81.3	65.6	100.0	90.6	100.0
d				93.8	59.4	68.8	93.8
e					90.6	25.0	81.3
f						84.4	53.1
g							93.8

[%]

Up FM		<i>N</i> = 8 participants					
	a	b	c	d	e	f	g
a	87.5	6.3	37.5	81.3	96.9	96.9	96.9
b		87.5	43.8	71.9	96.9	93.8	100.0
c			87.5	53.1	93.8	93.8	96.9
d				93.8	81.3	68.8	96.9
e					90.6	34.4	90.6
f						96.9	84.4
g							96.9

[%]

CF		<i>N</i> = 8 participants					
	a	b	c	d	e	f	g
a	96.9	15.6	12.5	43.8	93.8	81.3	90.6
b		93.8	12.5	31.3	62.5	87.5	87.5
c			93.8	37.5	84.4	90.6	87.5
d				100.0	59.4	43.8	87.5
e					96.9	25.0	43.8
f						96.9	18.8
g							96.9

[%]

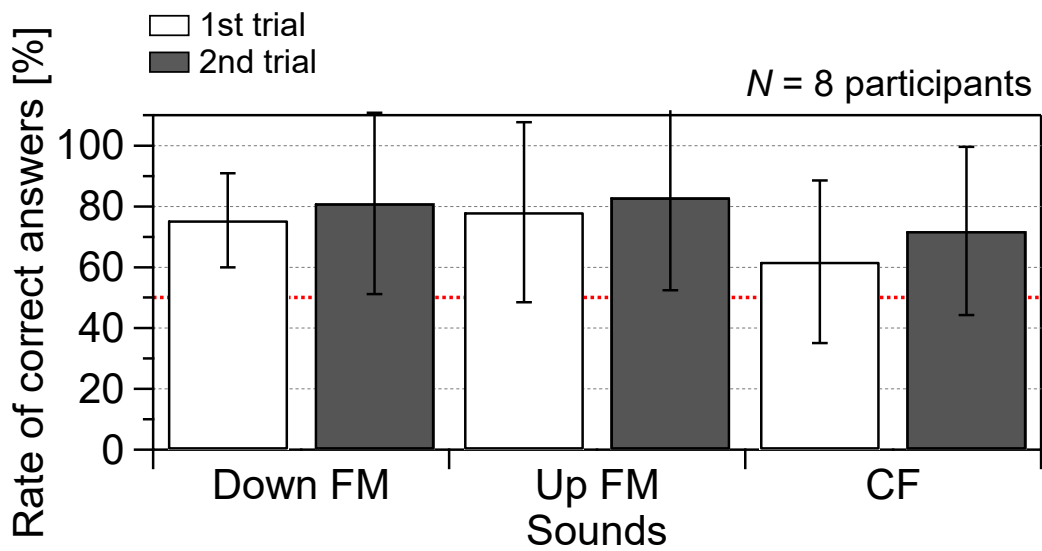


Figure 5.3 Average percent of correct answers performance of the eight participants under the downward, upward FM sweeps and CF burst signal conditions. The horizontal red dashed lines at the mean correct answer rate (50 %) indicate chance performance.

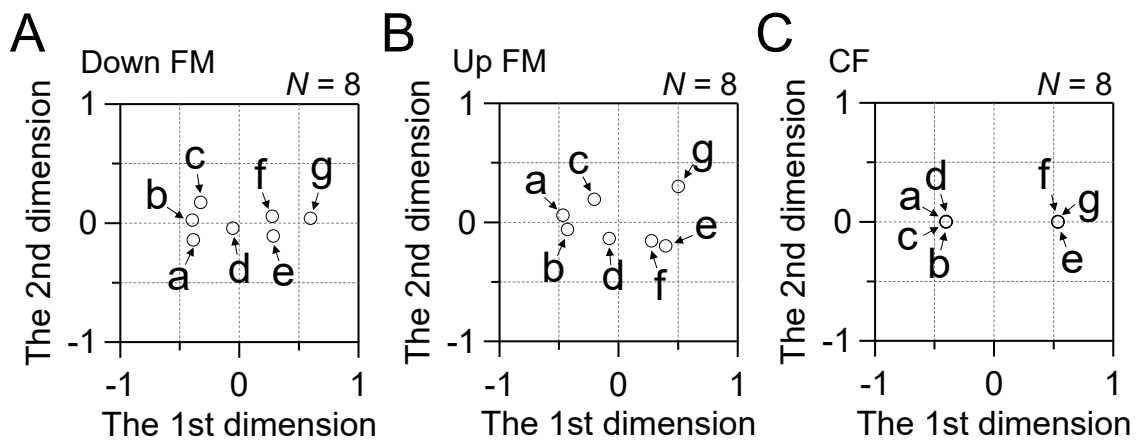


Figure 5.4 (A–C) Scatter plots of each relationship between the echoes from six different targets (a: acrylic board, b: wood, c: foamed polystyrene, d: artificial grass, e: artificial foliage, f: gardening fence) or one from no targets (g: no targets) based on the nonmetric multidimensional scaling method. In the numerical calculation, we used the averaged correct answer rate ($N = 8$) presented in Table 1 as the degree of unlikeness between targets. The values of the nonmetric stress criterion were 2.52×10^{-6} (A, Down FM), 7.03×10^{-7} (B, Up FM), and 2.54×10^{-6} (C, CF), respectively.

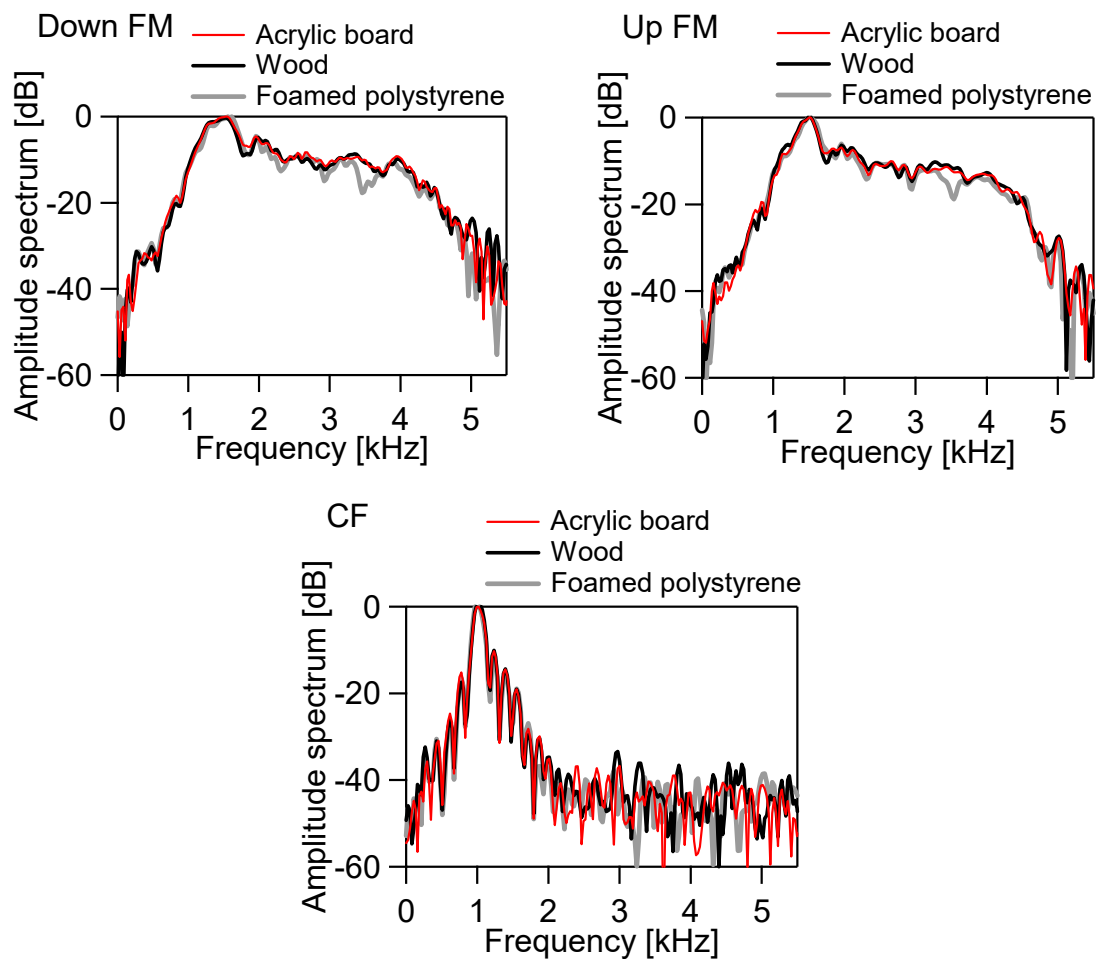


Figure 5.5 (A–C) Normalized amplitude spectra of echoes from the acrylic board (red line), wood (black line), and formed polystyrene (gray line) under downward (A) and upward (B) FM sweep conditions, and CF burst (C) condition, respectively.

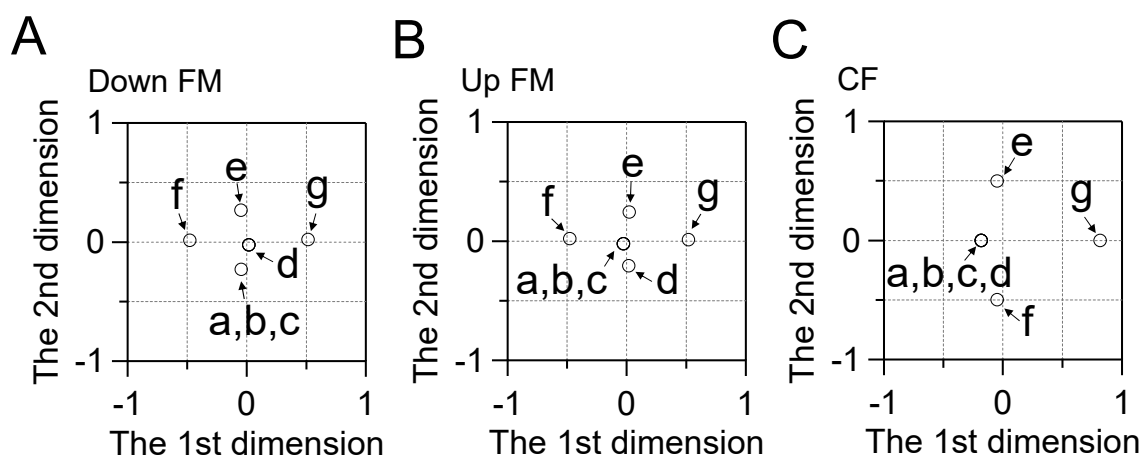


Figure 5.6 (A–C) Scatter plots of relationships among seven sounds under downward FM (A), upward FM (B), and CF (C) conditions. The values of the nonmetric stress criterion are 8.59×10^{-7} (A, Down FM), 6.81×10^{-7} (B, Up FM), and 5.99×10^{-7} (C, CF).

Table 5.2 Summary of GLMM statistics for the effects of the degree of similarity among the acoustic parameters (the peak value of the cross-correlation in the amplitude envelope, frequency spectra, and spectrogram) on the number of correct answers in each target pair ($N = 168$).

Sound	Fixed factor	Estimates	SE	z	p
Down FM	Amplitude spectrum	-0.009	0.001	-7.851	<0.001
	Amplitude envelope	-2.293	0.287	-7.982	<0.001
	Spectrogram	-0.057	0.007	-8.283	<0.001
Up FM	Amplitude spectrum	-0.008	0.001	-10.701	<0.001
	Amplitude envelope	-2.037	0.190	-10.717	<0.001
	Spectrogram	-0.057	0.005	-10.793	<0.001
CF	Amplitude spectrum	-0.008	0.001	-9.828	<0.001
	Amplitude envelope	-2.111	0.215	-9.813	<0.001
	Spectrogram	-0.046	0.005	-9.865	<0.001

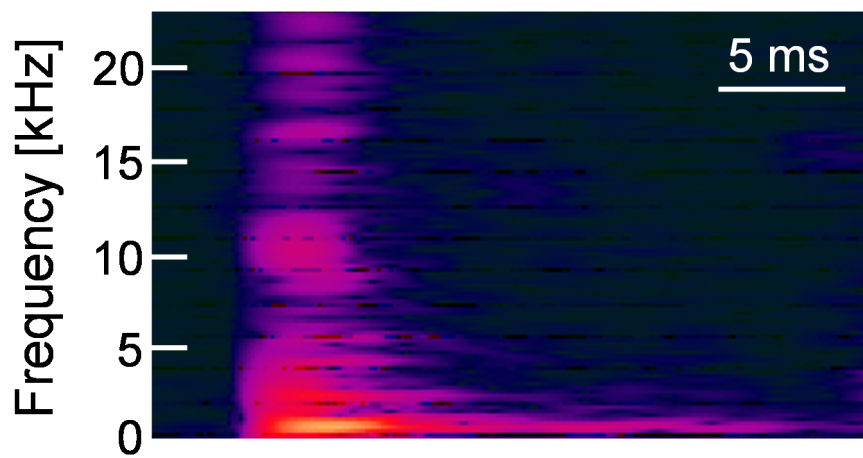


Figure 5.7 Spectrogram of click sound by a subject.

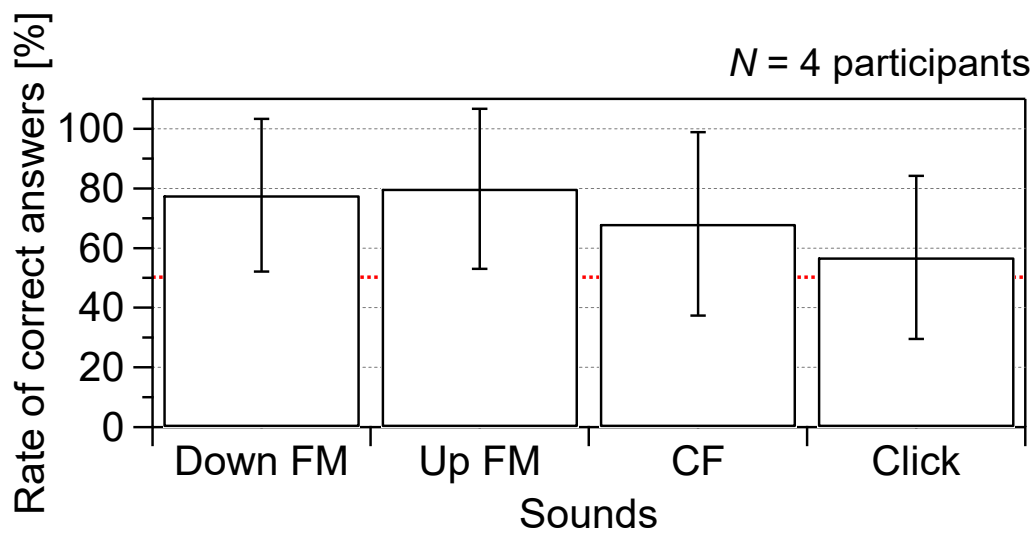


Figure 5.8 Averaged rate of correct answers of sound discrimination for two subjects under down FM, up FM, and CF conditions (ultrasonic binaural echo condition) and mouth click condition. The error bar represents standard deviation. The horizontal red dashed lines at the mean correct answer rate (50 %) indicate chance performance.

Chapter 6. Shape, Texture and Material Discriminations by Ultrasonic Binaural Echoes in Sighted Echolocation Novices

6.1 Introduction

Bats and dolphins emit ultrasounds and listen to the echoes from objects. This is termed as echolocation. Echolocation in the high-frequency range is an ideal sensory system to detect small objects, i.e., prey, and navigate complex obstacle spaces, because the sensing resolution depends on the wavelength of the echolocation sound. For example, echolocating Japanese house bats (*P. abramus*) use downward frequency-modulated (FM) ultrasonic pulses with harmonics, and the frequency of the fundamental component is exponentially modulated from 100 to 40 kHz (Hiryu *et al.*, 2007). The shortest wavelength corresponds to ~3 mm to detect and capture prey, including small hemipterans and dipterans in air (Hirai and Kimura, 2004). In addition, echolocating killer whales (*Orcinus orca*) use ultrasounds with the majority of the energy in the spectra between 20 and 60 kHz to detect and capture prey, such as Chinook salmon (*Oncorhynchus tshawytscha*) in water (Au *et al.*, 2004). Therefore, the wavelength of the echolocation signals essentially corresponds to the target size in bats and dolphins. This suggests that the bats and dolphins use echolocation sounds with a wavelength that is suitable to detect targets.

Echolocation is used to detect and localize targets, and to discriminate and identify targets in their environment. For example, echolocating Atlantic bottlenose porpoises (*Tursiops truncatus*) reliably discriminate between cylinders with different wall thicknesses or different materials (Hammer Jr and Au, 1980). It was suggested that the predominant cue used to discriminate among these cylinders is the time-separation pitch (TSP) (Hammer Jr and Au, 1980). However, echolocating bats can detect differences in hole depths smaller than 1 mm in a target by using differences in the spectral distribution of energy but not in the overall intensity (Simmons *et al.*, 1974). Echolocating *R. ferrumequinum* was reported to discriminate between insects using the characteristic amplitude- and frequency-modulation pattern by fluttering insects on returning echoes (von der Emde and Schnitzler, 1990). This is because a sufficient Doppler shift is caused by a high-frequency signal, such as the ultrasound used by echolocating bats. A recent

study examined the behavioral strategies for texture discrimination by echolocation in free-flying *E. fuscus* and suggested that the flying bat listens to changes in sound spectra from an echo to an echo to discriminate between objects (Falk *et al.*, 2011).

Moreover, in some cases, blind individuals use echolocation similar to bats and dolphins. However, echolocating blind individuals use mouth clicks as an echolocation sound because humans do not produce and listen to ultrasounds. Recently, Thaler *et al.* (2011) measured the brain activity pattern of early and late blind echolocation experts and suggested that the processing of click-echoes recruits brain regions that are typically devoted to vision as opposed to audition in both. The study was the first to measure brain activity of the echolocation expert during listening echoes and was from a neuroscience viewpoint. A recent review presents previous studies, ranging from classic to recent studies, involving psychophysical examinations (Thaler and Goodale, 2016). In these studies, the subjects actively emit mouth-clicks to determine the ability of perceiving location (azimuth), distance, and size of sound-reflecting surfaces (Thaler and Goodale, 2016). An example of a classic study reported that blind individuals who use echolocation can detect a difference of ~10 cm in depth at a distance of 60 cm (Kellogg, 1962). A recent study reveals that an individual who was blind from birth and learned to echolocate early in life could detect a change as low as 4° in the azimuthal position of a 150-cm tall pole in a two-interval two-alternative forced choice task (Thaler *et al.*, 2011). Furthermore, a recent study suggested that blind individuals who are experts in echolocation discriminate among objects based on their physical size (irrespective of their “acoustic size”), suggesting that size constancy may also operate during echolocation (Milne *et al.*, 2015a).

In contrast, long-term training is necessary to master echolocation using mouth clicks because it is difficult to generate mouth clicks with a high sound pressure level and to sufficiently listen to an echo from objects. Moreover, it is also increasingly difficult for echolocation novices to extract information from returning echoes as opposed to echolocation experts (Milne *et al.*, 2014; Milne *et al.*, 2015a). In order to overcome these issues, blind mobility aids using ultrasounds were developed based on echolocation systems, such as bats and dolphins (Ifukube *et al.*, 1991; Kay, 2000; Mihajlik *et al.*, 2001; Waters and Abulula, 2007; Sohl-Dickstein *et al.*, 2015). For example, a device developed by Ifukube *et al.* (1991) transmitted upswept or downswept FM ultrasounds in the range of 70–40 kHz or a constant frequency (CF) sound of 50 kHz for echolocation, and the

received signals by stereo recoding were time-stretched prior to presenting them to a user. This suggested that the FM sounds appeared superior to the CF sound in the perception of obstacles. Recently, Sohl-Dickstein *et al.* (2015) developed a device that transmits FM ultrasounds and binaurally receives echoes via artificial pinnae mimicked bats and then simply down-samples the signal and sends the down-sampled signal to the user via headphones. They demonstrated that the device can be effectively used to examine the environment and that the human auditory system rapidly adapts to these artificial echolocation cues. These studies suggest that these types of devices can potentially aid blind people in interacting with their environment.

In a previous study, we developed a new system for binaural recording in the ultrasonic range (Uchibori *et al.*, 2015). The system consists of a miniature dummy head (MDH) that is printed based on 3D shape data of a standard dummy head and a device that converts ultrasonic echoes into audible sounds (Uchibori *et al.*, 2015). The results indicated that the sounds captured by the proposed MDH system were more accurately localized by the participants and outside the head more than normal stereo sounds (Uchibori *et al.*, 2015). Therefore, we concluded that it may be useful to experience echolocation by MDH.

The main purpose of this study is the evaluation of the echolocation ability of sighted echolocation novices using ultrasonic echoes measured by MDH (i.e., ultrasonic binaural echoes). To achieve this, we examined the utility of ultrasounds in human echolocation based on the shape discrimination performance. Next, we conducted shape, texture, and material discrimination experiments using the ultrasonic binaural echoes that were measured by MDH and acoustic cues for discrimination. Hiryu *et al.* (2007) previously examined the biosonar strategy of bats during echolocation from an engineering viewpoint. We provide a discussion on the experimental results from the view point of a bat biosonar because this knowledge is applicable to human echolocation. We also included a discussion on the effective signal design for shape, texture, and material discriminations from an acoustics viewpoint to explore new possibilities of human echolocation using ultrasounds.

6.2 Materials and Methods

In this study, we conducted two types of discrimination experiments, namely Experiment 1 and Experiment 2. In Experiment 1, we examined if the downward FM ultrasound mimicking the echolocation sound of the bats was efficient in the discrimination of the 3D roundness of edge contours by the sighted echolocation novices. In Experiment 2, we examined if the sighted echolocation novices could echolocate by ultrasonic binaural echoes through shape, texture, and material discrimination experiments. Based on the discrimination performance, we examined the acoustic cues that were useful for discrimination by conducting a statistical analysis. In addition, we examined which sonar broadcast was effective for the shape, texture, and material discrimination in human echolocation in order to discriminate using the acoustic cue effectively. Here, the participants were required to discriminate echoes of targets that had differences in the roundness of the edge contours (shape), size of the cube on the object surfaces (texture), and categories of materials (material) in Experiment 2.

6.2.1 Participants

A total of 40 participants, P1–40 aged between 20 and 28 years (23.1 ± 1.8) participated in Experiments 1 and 2 at Doshisha University [Experiment 1, P1–10; Experiment 2, P11–20 (shape); P21–30 (texture); and P31–40 (material)]. Hearing tests revealed that pure tone thresholds of all participants were within a hearing level of 30 dB up to 5 kHz. All participants reported the absence of prior experience with echolocation and were found to possess normal or corrected normal vision. The testing was performed in a sound-attenuated chamber [2.3 m (H) \times 1.6 m (L) \times 1.4 m (W)]. The testing procedures were approved by the University ethics board and participants provided informed consent.

6.2.2 Targets

Figure 6.1 shows the 14 custom-made targets used for the shape (Shapes 1–5), texture (Textures 1–5), and material (Materials 1–4) discrimination experiments.

The Shapes 1–5 corresponded to the wooden poles (solid Narra-wood) with lengths of 80 cm and a smooth surface (Fig. 6.1A). As shown in Fig. 6.1A, Shapes 1–5

nearly appeared to be identical rectangles using vision as observed from the direction of the Z axis (Fig. 6.1A, bottom panel), depending on the shadows, and the five targets possessed different edge contours, as shown in the top panel in Fig. 6.1A. The corner radius increased from 0 to 4 cm in steps of 1 cm for Shapes 1–5, as shown in the bottom panel in Fig. 6.1A. Therefore, the lateral plane areas linearly decreased from 640 (Shape 1) to 0 cm² (Shape 5) with steps of 160 cm² for Shapes 1–5.

The Textures 1–5 were printed with a 3D printer (Replicator 2X Experimental 3D Printer; MakerBot, New York, USA) with acrylonitrile-butadiene-styrene (ABS) resin (Figs.1B). Textures 2–5 included multiple small cubes corresponding to 0.25 (Texture 2), 0.5 (Texture 3), 1 (Texture 4), and 2 cm (Texture 5) on a side on the base [8 cm (X) × 16 cm (Y) × 0.1 cm (Z)] (Fig. 1B). The cubes on the base were absent in Texture 1. The distance between the cubes was identical to the side of the cube, such that the total area of XY plane and area of the small cubes were identical (= 64 cm²), ensuring that the loudness of the echo essentially corresponded to the identical levels.

The Materials 1–4 were plates composed of four different materials, namely acrylic, glass, iron, and wood plates (Fig. 6.1C). The size of the Materials 1–4 was 15 cm (X) × 15 cm (Y) × 1 cm (Z).

6.2.3 Miniature dummy head

Figure 6.2 shows the 1/7 scaled MDH with a torso that was used to measure the ultrasonic binaural echoes from targets, composed of ABS resin using a 3D printer. We used 3D shape data of a standard dummy head (4128C Head and Torso Simulator, HATS; Brüel&Kjær, Nærum, Denmark) that was scanned using a 3D scanner (Model 2020i Desktop Laser Scanner; NextEngine, CA, USA) (Uchibori *et al.*, 2015). The MDH torso was formed such that it approximately imitated the shape of the HATS torso. The width, height, and depth of the MDH torso were 28, 51, and 23 mm, respectively. Omnidirectional condenser microphones with a diameter of 2.7 mm (B6 omnidirectional lavalier; Countryman Associates, Inc., California, USA) were inserted at the entrance of the left and right ear canals of the MDH. The interaural distance of the MDH was 22 mm.

6.2.4 Sound stimuli

Sound stimuli were created by convoluting the impulse response and arbitrary echolocation signal. The impulse response was measured by the upward logarithmic time stretched pulse (Log-TSP, Fig. 6.3) (Stan *et al.*, 2002) to eliminate harmonic distortion (Moriya and Kaneda, 2005). The length of the Log-TSP corresponded to 262,144 ($= 2^{18}$) points at a sampling rate of 192 kHz.

We measured the echo of the Log-TSP from each target in an anechoic chamber [3.6 m (W) \times 4.4 m (L) \times 3.4 m (H)] at Doshisha University using the experimental system shown in Fig. 6.4. The target was positioned at a distance of 15 cm from the loudspeaker (S-300HR; TEAC CORPORATION, Tokyo, Japan). The echoes from the target were measured by the two microphones with (binaural condition) and without (stereo condition) MDH, respectively. During the stereo recording, we measured the echoes under two conditions with microphone distances of 22 and 154 mm. The microphone distance of 22 mm was derived from the interaural distance of 1/7 scaled MDH. We also measured the echoes by monaural recordings without the target to measure the reference signal. The captured signals were digitalized with 32-bit accuracy at a sampling rate of 192 kHz using a high-speed USB audio interface (OCTA-CAPTURE; Roland, Shizuoka, Japan). The digitalized signals were synchronously stored with a personal computer (VAIO Z; VAIO Corporation, Nagano, Japan) using the audio editing software (Adobe Audition CC 2015). The sound pressure level of Log-TSP was 95 dB at a distance of 15 cm from the loudspeaker. The background noise level was approximately 14 dB (A), sufficiently low to measure the echoes from the targets.

The impulse response was calculated by a discrete Fourier transform after averaging the signals 32 times using MATLAB (MATLAB; The Math Works, Inc., Massachusetts, USA). The frequency characteristics of the loudspeaker and microphones were canceled by eliminating the reference signal measured by the monaural recording from the comparison signal measured by the stereo and binaural recordings.

There were two types of signal conditions; low- and high-frequency signal conditions (Fig. 6.5). Under the low-frequency signal condition, we used downward FM sweeps in a low-frequency range (1–5 kHz), shown in Fig. 6.5 (a), for the convolution operation with impulse response of each target that were measured in advance. We used a downward FM sweep, upward FM sweep, downward harmonic FM sweep, upward

harmonic FM sweep, band-limited noise, and CF bursts (7, 21, and 35 kHz) in a high-frequency range (7–35 kHz) shown in Figs. 6.5 (b–i) under the high-frequency signal condition.

In Experiment 1, the two downward FM sweeps in low- (1–5 kHz) and high- (7–35 kHz) frequency ranges were convoluted with the impulse responses measured by the stereo recordings at microphone distances of 154 (low-frequency) and 22 mm (high-frequency), respectively. The high-frequency convoluted signals (frequency band, 7–35 kHz; signal duration, 1 ms) were 1/7-times pitch converted to audible sounds (i.e., frequency band, 1–5 kHz; signal duration, 7 ms) by employing the time-expansion method using Adobe Audition CC 2015 to present the same for the participants. In contrast, the low-frequency convoluted signals (frequency band, 1–5 kHz; signal duration, 7 ms) were not pitch converted.

In Experiment 2, the signals ranging from 7–35 kHz were convoluted with the impulse responses measured by the binaural recording with the 1/7 scaled MDH. The high-frequency convoluted signals were 1/7-times pitch converted to audible sounds in the same way, as Experiment 1, before presenting for the participants. Each sound stimulus consisted of ten successive convoluted signals in a row with a time interval of 35 ms.

6.2.5 Procedure

Two different randomly-selected sound stimuli were presented in a two-alternative forced-choice task (2AFC) in Experiment 1 and in three-interval two-alternative forced-choice tasks (3I-2AFC) in Experiment 2 by using a custom-made program, EXPLAB (free software for computational experiment), through headphones (MDR-CD900ST, Sony, Tokyo, Japan). The time interval between the sound stimuli (i.e., inter-stimulus interval) was 300 ms. The loudness between the targets was not corrected. All sound stimuli were presented through headphones to the listeners below ~65 dB SPL at the position of the eardrum membrane by standardizing the maximum loudness in each signal condition based on ITU-R BS. 1770-3 loudness. In Experiment 1 the participants were required to accurately identify the sound stimulus that corresponded to echoes from the target with more rounded edge contours by immediately pressing the computer key upon identification. In Experiment 2, the participants were required to accurately identify

the sound stimulus that was identical to the second sound stimulus, the first or third, by immediately pressing the computer key upon identification.

We conducted four test sessions (without feedback of the answer on the display) after two or four training sessions (with feedback of the answer on the display) per participant. The participants for whom the percentage of correct answers was less than 75% in the first two training sessions were required to undergo an additional two training sessions. In the training sessions, the participants were required to discriminate between only two targets. Shapes 2 and 4 (shape discrimination experiment), Textures 2 and 4 (texture discrimination experiment), and Materials 2 and 4 (material discrimination experiment), were used in the training sessions of the shape, texture, and material discrimination experiments, respectively. In the test sessions, the participants were required to discriminate between five targets in each discrimination experiment (Shapes 1–5; Textures 1–5; and Materials 1–4). Different participants were involved in the shape (P11–20), texture (P21–30) and material (P31–40) experiments. There were a total of 28 trials (2 target pairs \times 2 signals \times 7 repetitions) and 40 trials (20 target pairs \times 2 signals \times 1 repetitions) in the training and test sessions in Experiment 1. In the shape and texture discrimination experiments in Experiment 2, there were a total of 96 trials (4 target pairs \times 8 signals \times 3 repetitions) and 320 trials (40 target pairs \times 8 signals \times 1 repetition) in the training and test sessions. In the material discrimination experiment in Experiment 2, there were a total of 96 trials (4 target pairs \times 8 signals \times 3 repetitions) and 192 trials (24 target pairs \times 8 signals \times 1 repetition) in the training and test sessions. After completing the experiment, a questionnaire survey on the acoustic cues (in Japanese) was conducted for each participant.

6.2.6 Statistics

A generalized linear mixed model (GLMM) (Schall, 1991) with a binomial distribution and logit link function was used to evaluate the following two points: the effect of the signal conditions (i.e., low and high frequency signal conditions) on the average percentage of correct answers and the effect of the degree of similarities in the acoustic parameters (i.e., amplitude wave pattern, frequency spectrum, spectrogram, and loudness) on the percentage of correct answers in each target pair. When multicollinearity was confirmed in the relationship between the acoustic parameters (the Pearson's

correlation coefficient exceeded 0.5), we excluded one of the acoustic parameters from the analysis. The participant's ID was treated as an explanatory variable describing a random effect. The number of the correct answers was treated as the response variable. We selected the model that led to the best estimation based on Akaike's Information Criterion (AIC) (Akaike, 1998). All analyses were performed using R Statistical Software Version 3.2.3. The analyses with GLMMs were conducted using the system function "glmmML".

6.3 Results

6.3.1 Discrimination of 3D edge contours by low- and high-frequency signals

In Experiment 1, we conducted the shape discrimination experiment by using low and high-frequency downward FM sweeps as shown in Fig. 6.5 (a, low-frequency downward FM sweep; b, high-frequency downward FM sweep). The participants were required to identify the sound stimulus that corresponded to the echoes from the target with more rounded edge contours.

As shown in Fig. 6.6, the average percentage of correct answers of all ten participants under the high-frequency signal condition was approximately 30% higher than that under the low-frequency signal condition (low-frequency, $59.0 \pm 9.7\%$; high-frequency, $88.6 \pm 7.3\%$). The GLMM evaluation indicated that the high frequency signal condition had a significant positive effect on the average percentage of correct answers ($\beta = 1.735 \pm 0.135$, $z = 12.851$, $p < 0.001$).

The frequency spectra of the impulse responses under the high-frequency signal condition exhibited a more complex spectral notch pattern than that under the low-frequency signal condition (Fig. 6.7). The differences among the five targets (Shapes 1–5) in the frequency spectra were also more significant under the high-frequency signal condition than under the low-frequency signal condition, suggesting that the impulse responses under the high-frequency signal condition contained more detailed 3D shape information. These results suggest that the high-frequency signals including ultrasound are more efficient for discriminating the roundness of edge contours than the low-frequency signals.

Table 6.1 shows that the performance of the participants was approximately 20%

more than the chance level (i.e., 50%) for any pair of targets under the high-frequency signal condition. Furthermore, some of the pairs including untrained targets (the pair of Shapes 1 and 4, Shapes 1 and 5, Shapes 2 and 5, and Shapes 3 and 5) had a higher percentage of correct answers than the training pairs (the pair of Shapes 2 and 4). These target pairs contained an echo that were not used in the trainings. This suggests that the participants were able to identify the sound stimulus that corresponded to echoes from the target with more rounded edge contours, even in the target pair that they listened to for the first time.

6.3.2 Effective signal design for discriminations of shape, texture and material

In Experiment 2, we conducted shape, texture, and material discrimination experiments by the 2AFC task for each of 10 participants (shape, P11–20; texture, P21–30; and material, P31–40) to examine the time-frequency structure of the signal that is efficient for these discriminations. The average percentage of correct answers are shown in Figs. 6.8A–C.

As shown in Figs. 6.8A and 6.8B, the participants performed well above the chance level under all signal conditions in the shape and texture discriminations. In the shape discrimination experiment, the average percentage of correct answers under all signal conditions was $71.4 \pm 2.7\%$ and it was higher than the chance level. The average percentage of correct answers under the band-limited noise condition was the highest (approximately 80%) of all.

In the texture discrimination experiment, the average percentage of correct answers under all signal conditions was $70.4 \pm 6.2\%$ and was higher than the chance level as the shape discrimination. The average percentage of correct answers under the CF (7 kHz) condition was the lowest, and the next lowest was the CF (21 kHz) condition. The average percentage of correct answers under the band-limited noise condition was the highest of all (~80%) of the texture discrimination experiments.

In the material discrimination experiment, the average percentage of correct answers under all signal conditions was $50.6 \pm 4.2\%$, similar to the chance level. Therefore, the material discrimination performance was approximately 20% lower than the shape and texture discriminations.

6.3.3 Frequency analysis of echoes

Figures 6.8D–F show the frequency spectra of the echo impulse responses of all targets. The echo impulse responses from Shapes 1–5 indicated differences in the levels of energy spectral density (ESD) among the targets while the notch patterns were almost identical. This implied that there were differences in the loudness of echoes that the participants listened to in the shape discrimination experiment among the targets. Additionally, the differences in levels of ESD among targets were slightly different based on the frequency band. This suggests that the spectral weight also differed among the targets.

Conversely, the echoes from Textures 1–5 exhibited different spectral notch patterns among the targets while the average differences of ESD in Textures 2–5 relative to Texture 1 were less than 2.5 dB (Fig. 6.8E). The spectral notches were observed in the different frequencies among the targets and the depths of spectral notches were also different among the targets, thereby suggesting that the timbre was different among Textures 1–5 while the loudness was almost identical.

In contrast, the echo impulse responses from the Materials 1–4 exhibited almost the same spectral pattern (Fig. 6.8F). This suggests that the material discrimination was more difficult than the shape and texture discriminations.

6.3.4 Acoustic cue

Table 6.2 shows the result of the GLMM for evaluating the effect of the shape, texture, and material discrimination performances and acoustic similarity of the echoes ($N = 100$ (shape), $N = 100$ (texture), and $N = 60$ (material) pairs for each discrimination experiment).

With respect to the shape discrimination, the differences of loudness among targets had a significant positive effect on the number of correct answers under each signal condition with the exception of the CF (35 kHz) burst condition (Table 6.2, top panel). This suggests that the participants used the loudness cue to discriminate shape.

Conversely, the differences in loudness did not exhibit any significant positive effects under any signal conditions in the texture discrimination (Table 6.2 middle panel). In contrast, the degree of similarity of frequency spectrum and/or spectrogram exhibited negative effects under the downward and upward FM sweep conditions. In addition, the

degree of similarity of envelope, frequency spectrum, and spectrogram exhibited negative effects under the downward and upward FM sweeps, the downward and upward harmonic FM sweeps, and the band-limited noise conditions. These suggested that the participants perceived differences in timbre among targets due to the differences among targets of the envelope, frequency spectrum, and spectrogram. Additionally, such significant negative effects were not observed under the CF burst conditions.

Finally, in the material discrimination, significant effects were not observed under other signal conditions under each signal condition with the exception of the CF (35 kHz) burst condition (Table 2, bottom panel). The degree of similarity of the envelope exhibited negative effects under the CF (35 kHz) burst condition.

6.4 Discussion

6.4.1 New possibility of human echolocation using high-frequency signal

From the experimental results in Experiment 1, we found that the average percentage of correct answers was above 80% when we used the high-frequency signal containing the ultrasound. This suggests that the participants were able to notice the acoustic characteristics that were changed consistently depending on the roundness of edge contours and then they could derive an answer. In addition, the participants were able to identify the sound stimulus that corresponded to echoes from the target with more rounded edge contours even in the target pair including the targets which were not used in the training. This is a very important ability in echolocation because it is necessary to sense an object that has never been sensed before.

The minimum wavelength of the signals used in this study (9 mm) was the almost same with the differences of the roundness of edge contours (10 mm step). Therefore, it is suggested that the participants could not judge which target had the rounder edge contours under the low-frequency signal condition because the minimum wavelength was approximately 60 mm, longer than the differences of the roundness of edge contours. Bats and dolphins use signals with a wavelength of echolocation signals essentially corresponding to the target size, suggesting that the bats and dolphins selectively use echolocation sounds with a wavelength suitable for detecting the targets. Therefore, it is important to select the frequency band depending on the targets, similar to bats and

dolphins.

Previous studies have demonstrated that the two-dimensional (2D) shapes of various objects, including circles, triangles, squares, or no targets could be discriminated by blindfolded sighted participants using self-generated mouth clicks (Rice, 1967; Hausfeld *et al.*, 1982; DeLong *et al.*, 2007b). These targets can be discriminated by vision. On the other hand, In Experiment 1, we used targets with almost the same rectangles when viewing from the *Z* axis direction (Fig. 6.1A), so we selected targets that were difficult to discriminate with vision, unlike the previous research. Using high-frequency signals, the subjects could discriminate the roundness of edge contours by listening to the echoes. The discrimination of edge contours is considered to be very important to sense the 3D shape because the “edge” would be the element forming the object shape. Therefore, our results suggest that echolocation allows us to “see” the 3D shape of objects which cannot be discriminated by visual information.

6.4.2 Robust acoustic cue for shape, texture and material discriminations

Differences in the spectral patterns of the echo impulse responses among the targets which were used in the shape, texture, and material discrimination experiments were contradistinctive. In concrete terms, the levels of ESD were different, whereas the spectral patterns were almost same among Shapes 1–5 (Fig. 6.8D). In contrast, the spectral pattern produced a different pattern whereas the levels of ESD were almost the same among Textures 1–5 (Fig. 6.8E). In addition, both the spectral pattern and the levels of ESD were almost the same among Materials 1–4 (Fig. 6.8F). Such a contradistinction among Shapes 1–5, Textures 1–5, and Materials 1–4 were consistent in the acoustic cue, as estimated by the GLMM analysis. According to the statistical evaluations by GLMM, it is suggested that the participants used the loudness cue in the shape discrimination while the timbre cue was used in the texture discrimination. In the material discrimination, there were no effective acoustic cue.

On the other hand, Milne *et al.* (2015a) demonstrated that the blind expert echolocators consistently identified the true physical size of the objects regardless of the target distance. This was interpreted as being due to the blind expert echolocators not using the loudness cue to identify the true physical size of the objects because the echo intensity changed depending on the target distance. Moreover, the loudness cue would

not be useful under actual environmental conditions because the loudness of the echoes may change by other factors in addition to the target distance, e.g., direction of the wind. Because the loudness cue would not be useful for the sensing in actual environmental conditions, we additionally conducted the discrimination experiments to investigate the shape discrimination performance without the loudness cue (the loudness level was corrected to the same level among the targets). The shape discrimination performance with (i.e., shape discrimination performance in Experiment 2, natural condition) and without loudness cue (loudness condition) are shown in Fig. 6.9A. As a result, the average percentage of correct answers under all signal conditions was $54.3 \pm 2.5\%$ under the loudness condition (Fig. 6.9A, open box) and was approximately 17% lower than that under the natural condition (Fig. 6.9A, filled box). However, an intriguing result was observed, especially for participant P13. Participant P13 performed better under the broadband signal conditions ($64.5 \pm 5.4\%$), compared to the performance under the CF burst conditions ($50.0 \pm 5.0\%$), and even the loudness levels were corrected to the same level among the targets (Fig. 6.9B). This suggests that it is possible to discriminate the 3D shape by using broadband signals in the high-frequency signals, even under the condition in which the loudness cue is not used. According to the questionnaire survey, participant P13 reported to use the timbre cue for the discrimination in the easy and difficult pairs, for both the natural and loudness conditions. Therefore, by imitating the sensing strategy of P13, it might be possible to perceive the roundness of the edge contours based on the timbre cue, even without the loudness cue. Additionally, in the participant P13 under the natural condition, the discrimination performance under downward FM sweep condition was higher than under upward FM sweep condition. A similar tendency was confirmed in the downward and upward harmonic FM sweep conditions under both natural and loudness conditions. As Ifukube *et al.* (1991), this might be because that faster high-frequency sound were not masked by later low-frequency sounds due to temporal masking.

Broadband signals are also suggested to be efficient for the texture discrimination. The participants were able to discriminate the texture by using the timbre cue under the broadband signal conditions, suggesting that even sighted echolocation novices may be able to discriminate the texture, regardless of the target distance, using the broadband signals (Fig. 6.8B). As for the CF sound, the correct answer rate of 35 kHz

CF sound was somewhat higher than the other CF sounds (7 and 21 kHz), probably because of the level difference occurring between targets due to the notch effect.

6.4.3 Comparison with the bat's ultrasonic sensing

In Experiment 2, we used signals that included signals that mimic bat echolocation sounds, assuming the echolocation occurs in the air. In addition to the time-frequency structure of the signal, the inter stimulus interval (35 ms) was also relatively mimicked with the bat's feeding buzz that are observed just before prey capture (Fujioka *et al.*, 2011). However, echolocating bats change the other acoustic parameters (e.g., pulse direction, terminal frequency, and sound pressure level) actively when flying in a 3D space (Fujioka *et al.*, 2011; Fujioka *et al.*, 2014; Motoi *et al.*, 2017). Therefore, it will be possible to improve the sensing performance further in human echolocation by examining the effect of those acoustic parameters which have not been taken into account in this study. However, the main difference from the actual echolocating bat was that the participants passively listened to the echo. Specifically, real bats navigate by active ultrasonic sensing while flying in a 3D space (Fujioka *et al.*, 2014). We have already developed software to convert ultrasound echoes into audible sounds in real time (Ashihara and Kiryu, 2008). Therefore, using this system, we aim to expand this to the echolocation experiment using active sensing in real time in future research. In addition, it is conceivable that ultrasonic binaural echoes measured by MDH become more operative in the perception of the moving targets than the fixed targets because the ultrasonic binaural echoes contain the 3D spatial information. Actually, we demonstrated that the sounds captured by the proposed MDH system were more accurately localized by the participants and outside the head more than normal stereo sounds (Uchibori *et al.*, 2015). Therefore, it is also required to extend the discrimination experiment using moving target or targets positioned at the various position in order to apply such an advantage of the ultrasonic binaural echoes.

In this study, we found that the material discrimination was more difficult than shape and texture discriminations for sighted echolocation novices (Fig. 6.8C). Nevertheless, echolocating *T. truncatus* was reported to discriminate cylinders of different materials by using TSP (Hammer Jr and Au, 1980). In addition, human subjects were reported to be able to discriminate materials using pitch converted to dolphin-like

signals (Au and Martin, 1989). Dolphins use broader ultrasonic pulses with shorter durations than bats, suggesting that they can acquire the impulse responses. Therefore, we additionally conducted a material discrimination experiment using pitch converted impulse responses of Materials 1–5. As a result, the average percentage of correct answers under the impulse response condition was not significantly higher, than that under the other signals. This suggests that it is necessary to reexamine signals design to discriminate material.

6.5 Summary

In this study, we examined if sighted echolocation novices could discriminate the 3D roundness of edge contours among five targets using the downward frequency-modulated (FM) ultrasound mimicking the echolocation sound of the bats. Our results showed that the participants could identify the roundness of edge contours by using downward FM echoes in the high-frequency range (7–35 kHz) that were converted to pitch in the audible range (1–5 kHz). In addition, the participants could discriminate between targets that were not used in the training. We additionally conducted shape, texture, and material discrimination experiments for other sighted echolocation novices to examine the suitable signal for the discriminations. Consequently, we found that the broadband signals were useful in shape, texture, and material discriminations, because timbre is a robust acoustic cue.

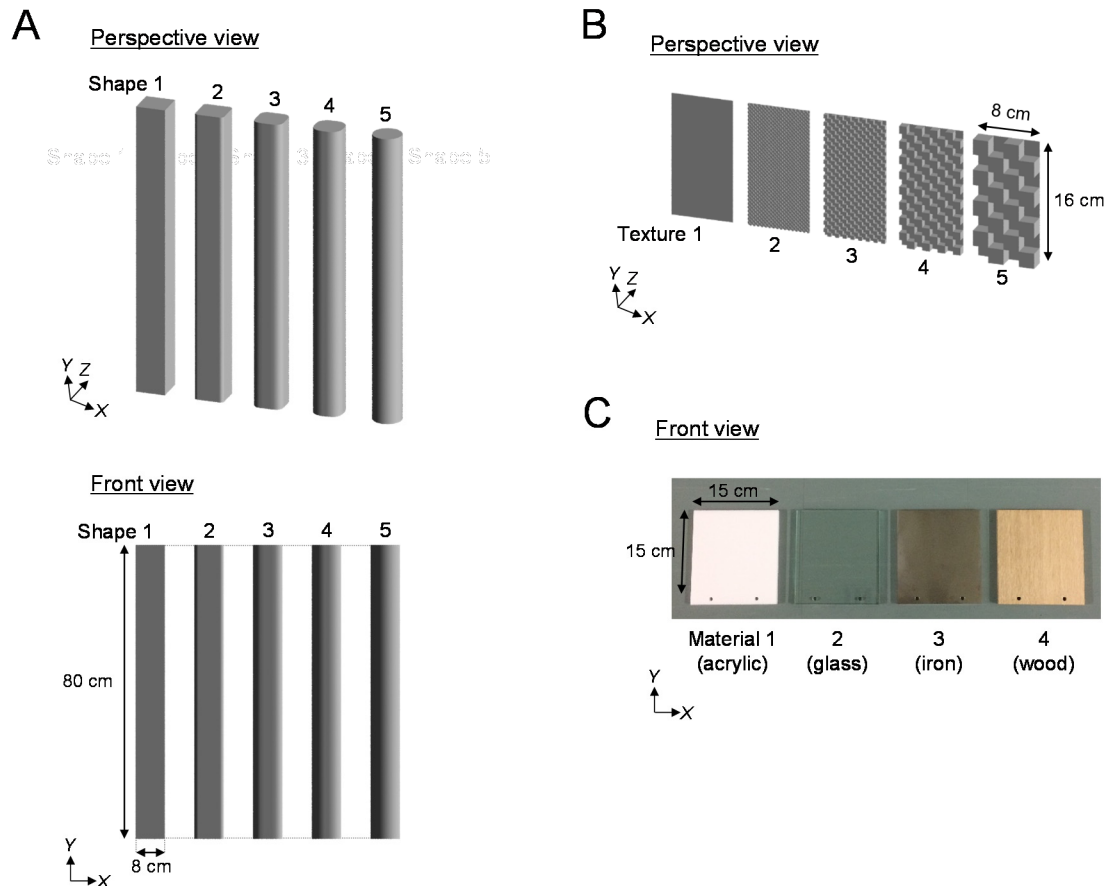


Figure 6.1 (A) Perspective (top panel) and front (bottom panel) views of schematic of five targets, Shapes 1–5 used in the shape discrimination experiments in Experiment 1 and 2. The Shapes 1–5 were made of Narra-wood. The length of the targets was 80 cm. The corner radius increases from 0 cm to 4 cm with a step of 1 cm for Shapes 1–5. The lateral plane areas linearly decreased from 640 (Shape 1) to 0 cm² (Shape 5) in steps of 160 cm² for Shapes 1–5. (B) Perspective view of the schematic diagram of the five targets, Textures 1–5 used in the texture discrimination experiment in Experiment 2. Textures 2–5 included multiple small cubes corresponding to 0.25 cm (Texture 2), 0.5 cm (Texture 3), 1 cm (Texture 4), 2 cm (Texture 5) on a side on the base [8 cm (X) × 16 cm (Y) × 0.1 cm (Z)]. Cubes on the base were absent in Texture 1. (C) Photograph of four targets, and the Materials 1–4 used in the material discrimination experiment in Experiment 2. The Materials 1–4 were plates composed of four different materials, namely acrylic (Material 1), glass (Material 2), iron (Material 3), and wood (Material 4) plates. The size of Materials 1–4 corresponded to 15 cm (X) × 15 cm (Y) × 1 cm (Z).

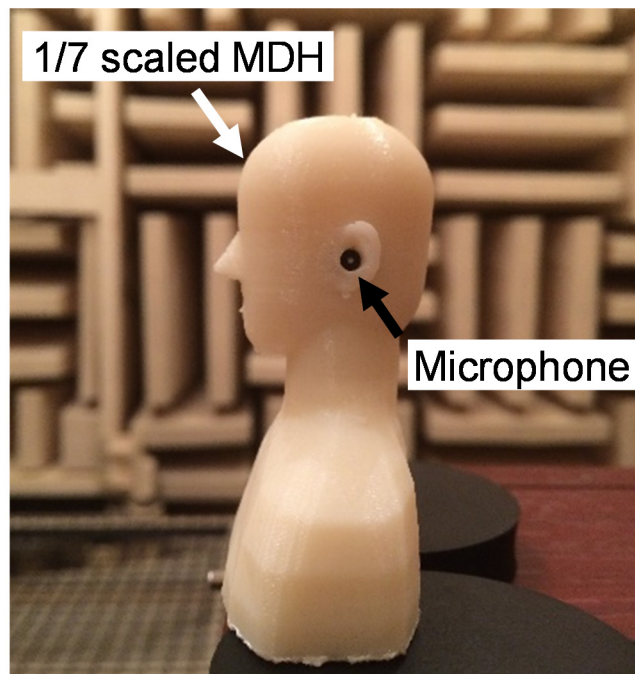


Figure 6.2 Photograph of the 1/7 scaled MDH which had two microphones inserted into the ear canals.

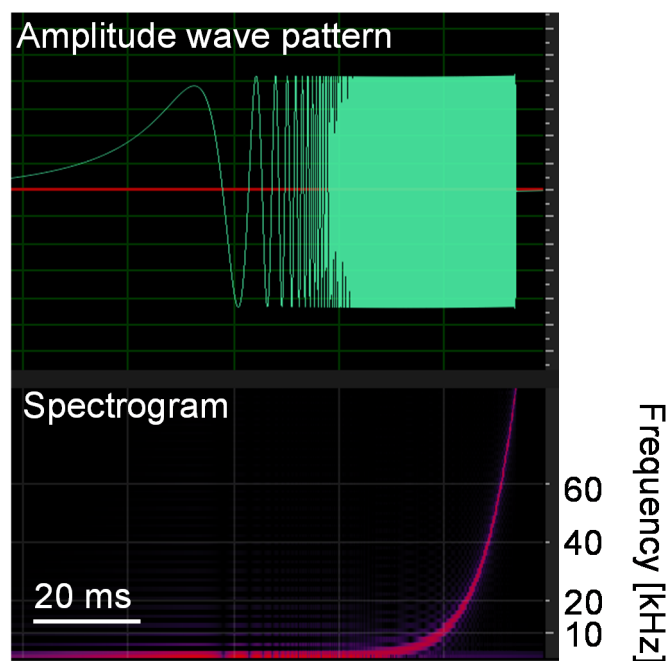


Figure 6.3 Amplitude wave form and spectrogram of the logarithmic sweep signal transmitted by the loud speaker.

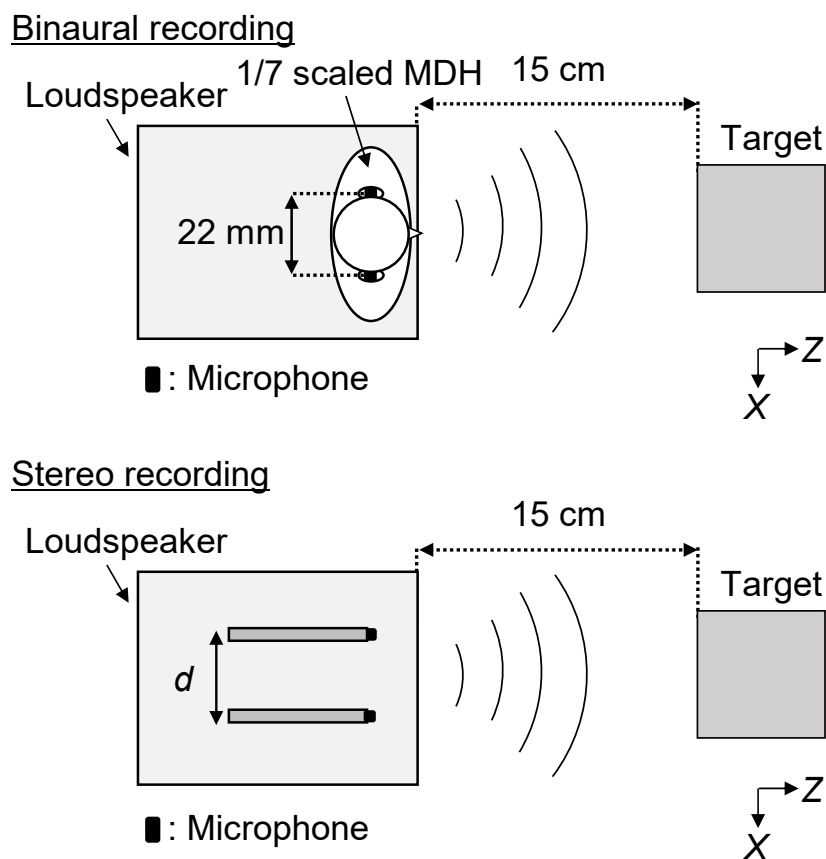


Figure 6.4 The schematic diagram of experimental system of the binaural (top panel) and stereo recordings (bottom panel).

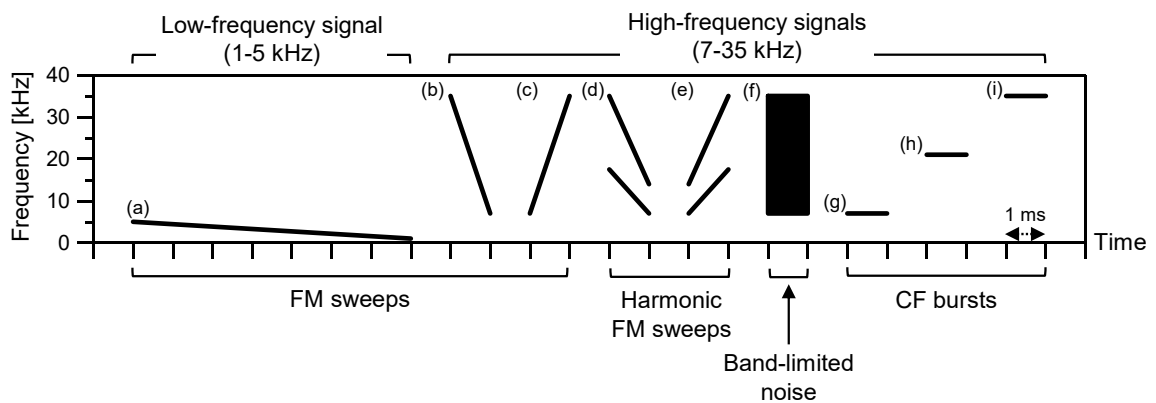


Figure 6.5 Spectrograms of nine signals (a, downward FM sweep (1-5 kHz); b, downward FM sweep (7-35 kHz); c, upward FM sweep (7-35 kHz); d, downward harmonic FM sweep (7-35 kHz); e, upward harmonic FM sweep (7-35 kHz); f, band-limited noise (7-35 kHz); and g–i CF bursts (7, 21, 35 kHz) which were convoluted with an impulse response measured in advance.

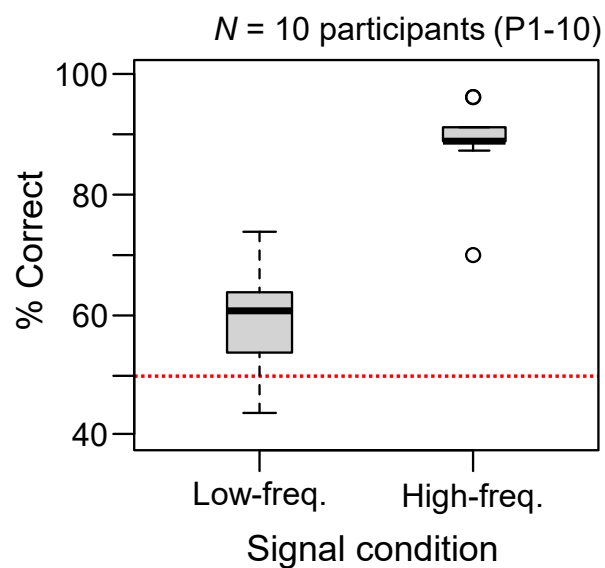


Figure 6.6 Average percentage of correct performance of the ten participants P1–10 in the low- and high-frequency signal conditions. The horizontal red dashed lines at the mean correct answer rate = 50% indicate a chance performance.

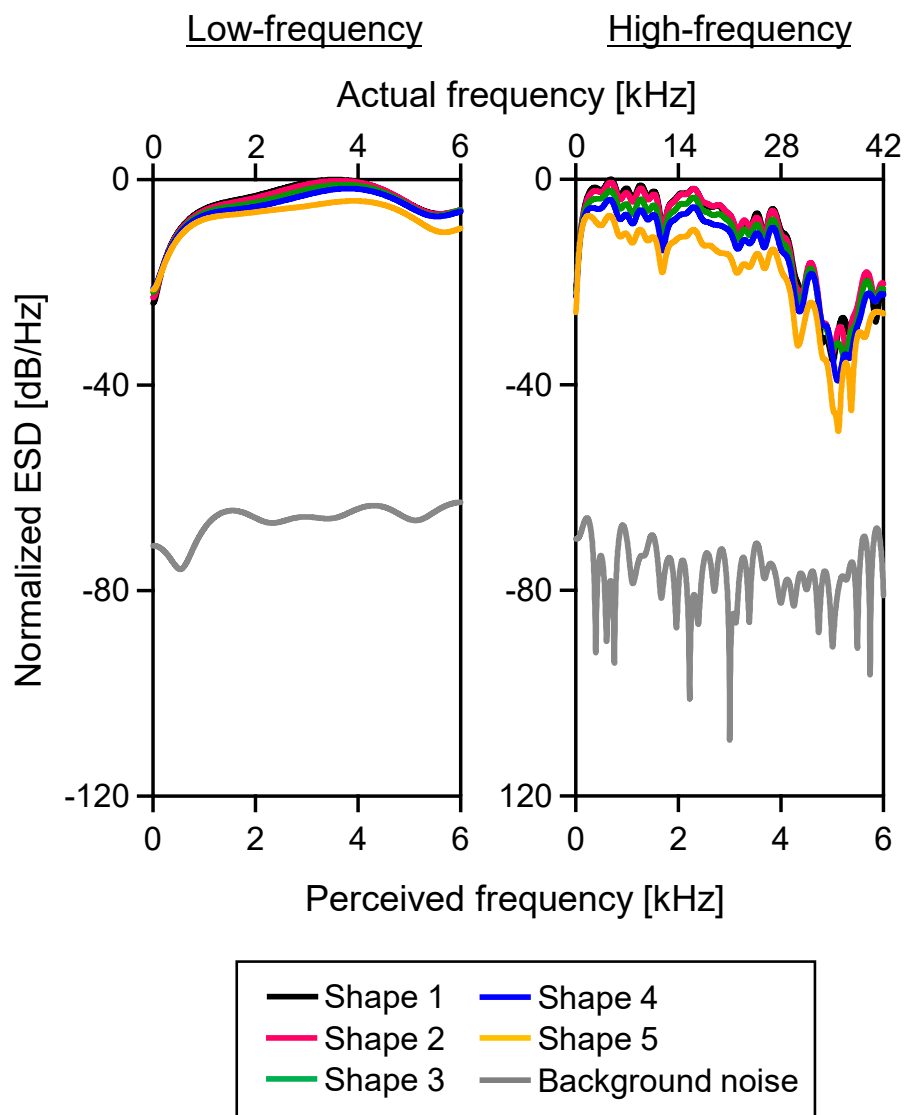


Figure 6.7 Frequency spectra (energy spectral density) the impulse response of Shapes 1–5 (Shape 1, black line; Shape 2, magenta gray line; Shape 3, green line; Shape 4, blue line; and Shape 5, yellow line) which were convolved with the low- (left panel) and high-frequency (right panel) downward FM sweeps. Frequency spectra of background noise are shown as the gray lines. The top and bottom horizontal axes show the actual and perceived (pitch converted) frequencies, respectively.

Table 6.1 Average percentage of correct answers in each target pair of the 10 participants (P1–10) under the high-frequency signal condition in Experiment 1.

High-frequency		<i>N</i> = 10 participants			
	Shape 1	Shape 2	Shape 3	Shape 4	Shape 5
Shape 1		68.8	86.3	95.0	93.8
Shape 2			81.3	90.0	96.3
Shape 3				88.8	97.5
Shape 4					88.8
Shape 5					

[%]

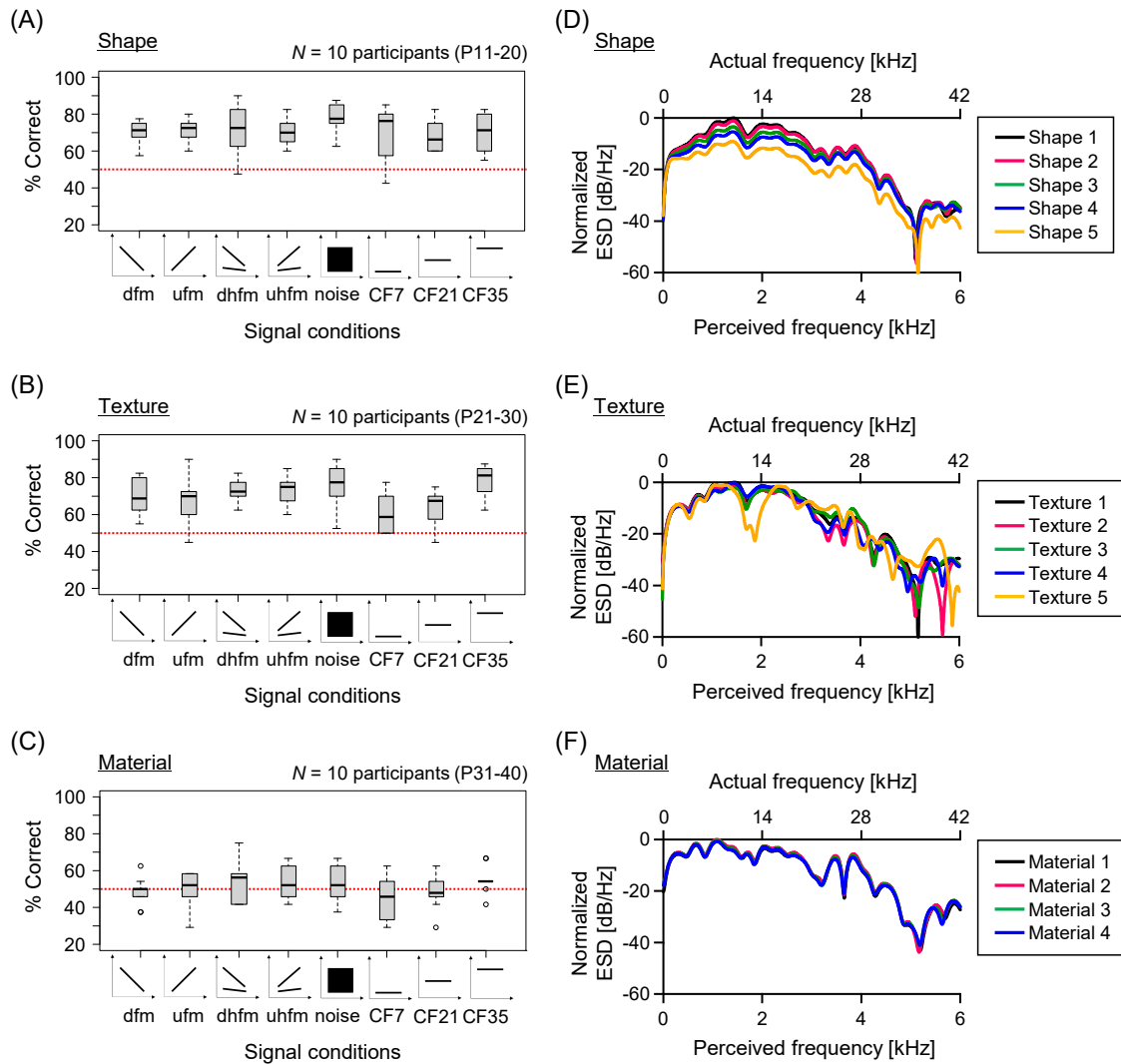


Figure 6.8 (A-C) Average percentage of correct answers performance of the 30 participants (P11–40), under eight signal conditions in the shape (A), texture (B), and material (C) discrimination experiments, using ultrasonic binaural echoes. The horizontal red dashed lines at mean correct answer rate = 50% indicate chance performance. (D-F) Frequency spectra (energy spectral density) of the echo parts of the impulse responses of the Shapes 1–5 (D), Textures 1–5 (E) and Materials 1–4 (F). The pitch was converted to 1/7 times by time expansion method. The top and bottom horizontal axes show the actual and perceived (pitch converted) frequencies, respectively.

Table 6.2 Summary of GLMM statistics for the effects of the degree of similarity among the acoustic parameters (the peak value of the cross-correlation in the amplitude envelope, frequency spectra, and spectrogram) and the absolute value of the differences in the loudness on the number of correct answer in each target pair in the shape (top panel, $N = 100$), texture (middle panel, $N = 100$), and material (bottom panel, $N = 60$) discrimination experiments. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Shape						$N = 100$
Signal condition	Fixed factor	Estimates	SE	z	p	
Down FM sweep	loudness	0.292	0.051	5.789	<0.001***	
Up FM sweep	loudness	0.267	0.051	5.226	<0.001***	
Down harmonic FM sweep	loudness	0.249	0.052	4.790	<0.001***	
Up harmonic FM sweep	loudness	0.333	0.054	6.136	<0.001***	
Band-limited noise	loudness	0.309	0.058	5.283	<0.001***	
CF burst (7 kHz)	loudness	0.286	0.060	4.794	<0.001***	
CF burst (21 kHz)	loudness	0.435	0.057	7.601	<0.001***	
CF burst (35 kHz)	envelope	-29.639	4.806	-6.167	<0.001***	

Texture						$N = 100$
Signal condition	Fixed factor	Estimates	SE	z	p	
Down FM sweep	frequency spectrum	-0.051	0.022	-2.311	0.021*	
	spectrogram	-0.008	0.003	-2.600	0.009**	
Up FM sweep	frequency spectrum	-0.111	0.024	-4.636	<0.001***	
	loudness	-0.549	0.224	-2.449	0.014*	
Down harmonic FM sweep	envelope	-10.416	2.034	-5.120	<0.001***	
Up harmonic FM sweep	envelope	-9.344	2.074	-4.506	<0.001***	
Band-limited noise	envelope	-11.134	4.036	-2.759	0.006**	
	loudness	-0.666	0.204	-3.271	0.001**	
CF burst (7 kHz)	frequency spectrum	-0.001	0.001	-1.580	0.114	
	loudness	-0.420	0.168	-2.495	0.013*	
CF burst (21 kHz)	envelope	3.249	2.332	1.394	0.163	
CF burst (35 kHz)	spectrogram	0.021	0.014	1.461	0.144	

Material						$N = 60$
Signal condition	Fixed factor	Estimates	SE	z	p	
Down FM sweep	frequency spectrum	-0.068	0.061	-1.126	0.260	
Up FM sweep	loudness	0.463	0.551	0.839	0.402	
Down harmonic FM sweep	frequency spectrum	-0.051	0.035	-1.469	0.142	
Up harmonic FM sweep	loudness	1.492	0.573	2.603	0.009**	
Band-limited noise	loudness	-1.285	0.583	-2.206	0.027*	
CF burst (7 kHz)	loudness	0.000	0.000	0.913	0.361	
CF burst (21 kHz)	spectrogram	-0.010	0.017	-0.573	0.567	
CF burst (35 kHz)	envelope	-31.538	11.064	-2.850	0.004**	

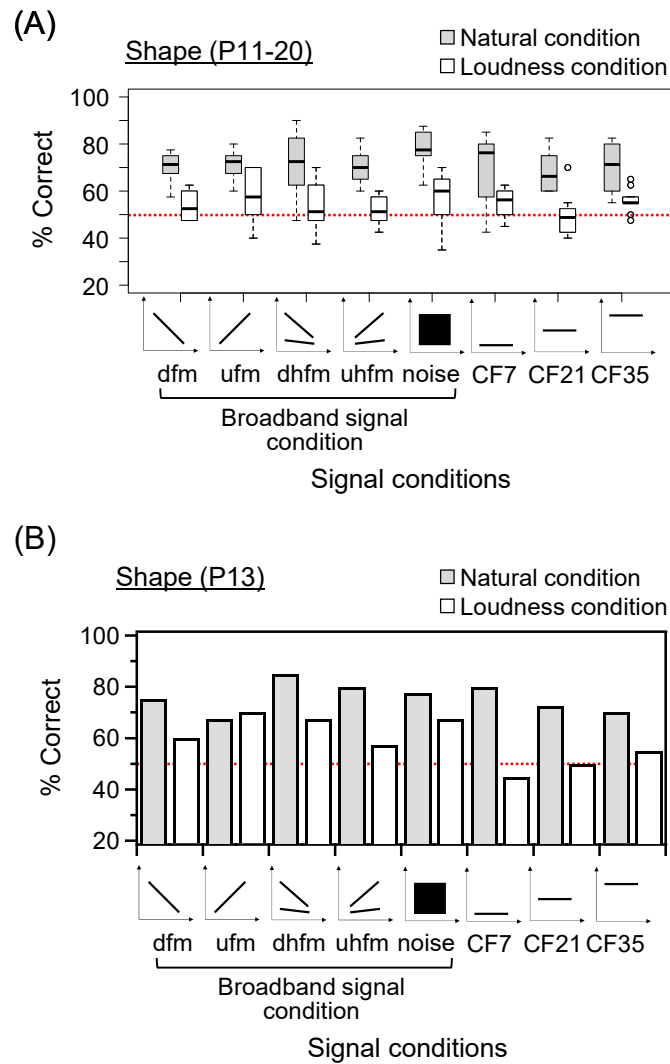


Figure 6.9 (A) Average percentage of correct performance of the ten participants P11-20 in the natural (loudness was not corrected, i.e., Experiment 2) and loudness (loudness was corrected to the same level) conditions. The horizontal dashed lines at the mean correct answer rate = 50% indicate a chance performance. (B) Average percentage of correct performance of the participant P13 with absolute hearing under eight signal conditions in the shape discrimination experiment. The filled and open bars show the average percentage of the correct answers performance under the natural and loudness conditions, respectively. The horizontal dashed lines at the mean correct answer rate = 50% indicate a chance performance.

Chapter 7. Conclusions

In this dissertation, the unique and sophisticated sonar strategy of wild echolocating bats were examined based on both field measurement and mathematical modeling. In addition, we successfully conducted original psychoacoustic experiments for sighted echolocation novices by determining acoustic parameters of echolocation signals (e.g., time-frequency structure, inter-pulse interval) based on the knowledge acquired in the field measurement. Furthermore, we studied acoustic cues for target discrimination by comparison between experimental results and acoustic analysis of echoes; a new acoustic sensing method utilizing the advantages of ultrasound was proposed. In this chapter, we summarize the main results of the dissertation and discuss directions for future work.

7.1 Summary of Main Results

7.1.1 Coordinated control of the acoustical field of view and flight in three-dimensional space for consecutive capture by echolocating bats during natural foraging (Chapter 3)

The direction and directivity pattern of the ultrasound broadcast of these bats are important factors that affect their acoustical field of view. These allow us to investigate how the bats control their acoustic attention (pulse direction) for advanced flight maneuvers. The purpose of this study was to understand the behavioral strategies of acoustical sensing of wild Japanese house bats (*P. abramus*) in 3D space during consecutive capture flights. The results showed that the bats successively captured multiple airborne insects in short time intervals (less than 1.5 s); they maintained both the immediate and subsequent prey simultaneously within the beam widths of the emitted pulses in both horizontal and vertical planes before capturing the immediate one (Sumiya *et al.*, 2017). We numerically simulated the bats' flight trajectories when they approached two prey items successively to investigate the relationship between the acoustical field of view and the prey direction for effective consecutive captures. This simulation demonstrated that acoustically viewing both the immediate and subsequent prey simultaneously increases the success rate of capturing both prey items, which is considered to be one of the basic axes of efficient route planning for consecutive capture

flight. Our findings suggest that the bats then keep future targets within their acoustical field of view for effective foraging. In addition, in both the experimental results and the numerical simulations, the acoustic sensing and flights of the bats showed narrower vertical than horizontal ranges. This suggests that bats control their acoustic sensing according to different schemes in the horizontal and vertical planes depending on their surroundings. These findings suggest that echolocating bats coordinate their control of the acoustical field of view and flight for consecutive captures in 3D space during natural foraging.

7.1.2 Mathematical modeling of flight and acoustic dynamics of an echolocating bat during multiple-prey pursuit (Chapter 4)

We proposed a new mathematical model describing the nonlinear dynamics of the flight and pulse directions of an echolocating bat approaching two successive targets (Sumiya *et al.*, 2015). In the model, a bat is assumed to control its flight and pulse directions depending on the directions of the targets. Numerical simulation of the present model shows that the model bat successfully captures both targets within a short time interval without losing them from its sonar beam at specific parameter values. The simulation also suggests that the successive prey capture is completed when the echolocation pulses are directed to the subsequent target before capturing the immediate target. Such a relationship between the flight and acoustic sensing can be also observed in the behavioral data of wild bats.

Our numerical simulation demonstrated that the present model can qualitatively explain successive prey capture with specific parameter values. In addition, it is suggested that such successive prey capture is accomplished when a model bat flies toward the immediate target while emitting pulses toward the subsequent target.

7.1.3 Human echolocation by ultrasonic binaural echoes (Chapter 5)

In chapter 5, we showed research results on the establishment of a human echolocation system that takes advantage of bat echolocation. Based on the proposed human echolocation system, we investigated the discrimination ability of sighted subjects for object texture to understand acoustic cues for texture recognition in human echolocation. FM and CF ultrasonic echoes from six objects with different materials and

surface structures were acquired using a 1/7-size miniature dummy head for presentation of 1/7-times pitch converted binaural audible sounds to listeners through headphones. The average rate of correct answers in the case of extremely different surface conditions (i.e., acrylic board vs. artificial foliage) was more than 90%; in the slightly different surface conditions (i.e., acrylic board vs. foamed polystyrene), it was under 40%. These suggest that it is possible to discriminate targets with extremely different surface structures and materials, which are easy to discriminate by vision, by listening to ultrasonic binaural echoes. Furthermore, the rate of correct answers in the CF sound condition was approximately 13% lower than those in the FM sound condition. The correlation diagram among targets by multidimensional scaling was dispersed more remarkably in the FM sound condition. When the target pair had slightly different surface conditions, differences in the notch pattern of amplitude spectra were observed especially in the FM sound condition. These suggest that FM ultrasonic binaural sound is more effective for slightly different texture perception than CF ultrasonic binaural sound. In addition, we found that ultrasonic binaural echoes might be more effective for human echolocation than mouth clicks through a comparison of discrimination ability.

7.1.4 Shape, texture, and material discriminations by ultrasonic binaural echoes in sighted echolocation novices (Chapter 6)

To evaluate the utility of ultrasound for target discrimination, we examined if sighted echolocation novices could determine the 3D roundness of edge contours among five targets using the downward FM ultrasound that mimics the echolocation sound of the bats. These targets were difficult to discriminate by vision without shadows. Our results showed that the participants could identify the roundness of edge contours using downward FM echoes in the high-frequency range (7–35 kHz) that were converted to pitch in the audible range (1–5 kHz). In addition, the participants could discriminate between targets that were not used in the training. We also conducted shape, texture, and material discrimination experiments for other sighted echolocation novices to examine the suitable signal for the discriminations. We found that the average percentage of correct answers in shape and texture discriminations were both approximately 70%, and the average correct answer rate for material discrimination was approximately 50%. From the frequency analysis of the echo and the statistical analysis using the generalized linear

mixed model, it was found that the participants were able to discriminate shape and texture using loudness (shape) and timber (texture) cues, respectively. In material discrimination, the average percentage of correct answers was approximately equal to the chance level (=50%) because there were no effective clues. Moreover, when we performed the shape discrimination experiment with the loudness cue removed, the average percentage of correct answers of all signals decreased by approximately 17%. However, in some subjects, the average percentage of correct answers was about 10% higher than the chance level when the broadband signal was used even when the loudness cue was removed (the average percentage of correct answers under CF sound is less than the chance level). Based on this, it was suggested that even in situations where loudness cannot be used as an acoustic cue, it could be used as a supplemented to the timber cue.

7.2 Future studies

7.2.1 Background, issue, and resolution

If we can elucidate the sensory perception of “see by sound” in the future, it could lead to the proposal of a new acoustical sensing technology to identify and navigate obstacles by sound rather than vision like in bats and dolphins. In addition, we believe that the sensory perception of “seeing by sound” that we pursued in this research has applications in cross-modal sensory information presentation technologies such as virtual reality (VR) and augmented reality (AR). The possible applications of this research are wide, e.g., new sensing technology as entertainment or guidance system for humans using sound. Furthermore, a new breakthrough can be expected for support technology for the visually handicapped. Urgent countermeasures are required because blind people have fallen from the platforms or have been killed by trains; accidents occur frequently for them. Therefore, in future studies, we aim to develop a next-generation acoustic sensing technology that imitates the active sensing of bats for humans to have the ability to “see by sound.” This is a new field of biomimetics that uses mechanisms obtained from living things to humans instead of technology. Specifically, we will find a sensing method that can efficiently and actively extract information according to the user's purpose (e.g., distance measurement, direction localization, shape identification, material identification). We will perform psychoacoustic experiments and brain activity

measurements for echolocation using ultrasounds with high spatial resolution. Furthermore, we will try to develop wearable echolocation devices that can sense real environment using echolocation in a real time, and we aim to realize navigation using echolocation (Here, we define navigation as efficiently moving to destination.) using such a device. In this way, we try to establish the creation of human echolocation research aiming at application to the next generation of engineering technology, which leads from basic research to engineering application. In the past, we have learned the importance of conducting research from a multilateral viewpoint while proactively conducting research while crossing more than one discipline. From that experience, we conceived a challenging study which try to basic to engineering applications by collaborating of psychology, brain science and engineering.

7.2.2 Research content

In future studies, we will explore a sensing method that can efficiently and actively extract information according to the user's purpose (e.g., distance measurement, direction localization, shape identification, material identification). We will study echolocation using sound with suitable wavelength through psychoacoustic experiment and brain activity measurements. In addition, we will develop a wearable echolocation device that can transmit them in the real environment and real time. We aim to be able to navigate (move efficiently to the destination) using echolocation.

To accomplish the goal, we will perform the following three tasks: (1) psychophysical evaluation of object perception by echolocation, (2) measurement of brain activity during the echolocation task, and (3) development of wearable echolocation device based on the active sensing of bats.

In task (1), we will first perform psychoacoustic experiments on the perception of distance, direction, shape, and material, which are necessary information for "object identification," and which was previously untouched. We will also clarify the necessary acoustic cues to perceive each object information. Furthermore, we will construct a sensing system that allows the participants to actively change the acoustic parameters (e.g., frequency, pulse duration, sound pressure level, pulse emission rate) of the transmitted signal, and examine more effective sensing methods using intuition and potential capacity of the participants efficiently. We will also compare these results with

behavioral findings obtained by field measurements of sonar behavior of bats and establish a sensing strategy in human echolocation based on scientific knowledge.

In task (2), to evaluate the sensing method proposed in task (1) based on objective indicators, the brain activities of participants during echolocation will be analyzed using functional magnetic resonance imaging (fMRI).

In task (3), to apply the findings obtained in tasks (1) and (2) to navigation, we will develop a wearable echolocation device that can use echolocation while moving in a real environment and real time using the active sensing system constructed in task (1) as the foundation. We will also conduct experiments on navigation behaviors using the developed devices.

7.2.3 Features and originality of the research

In this research, we will examine a sensing method and a training method that can extract information more efficiently through acoustic psychological experiments based on active sensing of bats and brain activity measurement. Furthermore, to apply findings acquired through experiments to navigation, we will develop a wearable echolocation device that can sense in real environment and real time. Thus, the features and originality of the proposed research is that a researcher with experience in biosonar research in bats will try to from basic research to engineering applications by crossing multiple fields of psychology, neuroscience, and engineering.

Researchers specializing in psychology and neuroscience have mainly measured the echolocation capability of blind expert echolocators who use echolocation techniques in their lives. However, there have been few cases that biosonar researchers worked directly to analyze acoustic psychological experiments based on an acoustic viewpoint and analyze acoustic features of echoes as acoustic cues for object discrimination (DeLong *et al.*, 2007b; Wallmeier and Wiegrebe, 2014a; c). In addition, although some assistive devices for blind people have been developed (Ifukube *et al.*, 1991; Sohl-Dickstein *et al.*, 2015), there have been no studies to actively apply cross-modal sensory information presentation technology as AR and VR for all people.

Moreover, if we can confirm the brain activity related to vision and tactile sense even at echolocation in task (2) of the future research, it might be possible to establish a completely new training method of echolocation by positively inducing that brain activity

using neurofeedback technology in the future. Since this will also lead to support for blind individuals, it can be expected to have a great impact on welfare engineering and rehabilitation science. By completing this research, we can pioneer the possibility of new biomimetics that applies useful mechanisms obtained from animals to human. We believe that we can develop a field of new research that not only bats but also unique sensing ability that various organisms have refined for living are applied to human, in a modern society that artificial intelligence are about to go beyond human ability.

7.3 Final remarks

Using original, highly-developed, large-scale 3D microphone-array system and mathematical modeling, we studied the coordinated control of acoustical field of view and flight in three-dimensional space for consecutive capture by echolocating bats during natural foraging. In addition, in the human echolocation study conducted in parallel with the biosonar research in wild bats, we proposed the possibility of a new acoustic sensing method using ultrasounds; this is based on the result of the psychoacoustic experiment using proposed ultrasonic binaural echoes measured by MDH. In particular, our results showed that the participants able to perceive the roundness of edge contours, especially using ultrasound, which is difficult to distinguish visually without information of shadows. This suggests that echolocation using ultrasound not only becomes a substitute for the visual sense but may also be able to have more functions. Based on our field research in echolocating bats and our new human echolocation study, we can find new possibilities for human echolocation using ultrasound by performing the tasks listed in future studies.

References

- Aihara, I., Fujioka, E., and Hiryu, S. (2013). "Qualitative and quantitative analyses of the echolocation strategies of bats on the basis of mathematical modelling and laboratory experiments," *PloS one* **8**, e68635.
- Akaike, H. (1998). "Information theory and an extension of the maximum likelihood principle," in *Selected Papers of Hirotugu Akaike* (Springer), pp. 199-213.
- Arias, C., and Ramos, O. A. (1997). "Psychoacoustic tests for the study of human echolocation ability," *Applied Acoustics* **51**, 399-419.
- Arnott, S. R., Thaler, L., Milne, J. L., Kish, D., and Goodale, M. A. (2013). "Shape-specific activation of occipital cortex in an early blind echolocation expert," *Neuropsychologia* **51**, 938-949.
- Ashihara, K., and Kiryu, S. (2008). "Audio pitch shift using Hilbert transform," *IEICE Technical Report EA* **108**, 51-56.
- Au, W. W., and Herzing, D. L. (2003). "Echolocation signals of wild Atlantic spotted dolphin (*Stenella frontalis*)," *The Journal of the Acoustical Society of America* **113**, 598-604.
- Au, W. W., and Martin, D. W. (1989). "Insights into dolphin sonar discrimination capabilities from human listening experiments," *The Journal of the Acoustical Society of America* **86**, 1662-1670.
- Au, W. W. L., Ford, J. K. B., Horne, J. K., and Allman, K. A. N. (2004). "Echolocation signals of free-ranging killer whales (*Orcinus orca*) and modeling of foraging for chinook salmon (*Oncorhynchus tshawytscha*)," *The Journal of the Acoustical Society of America* **115**, 901-909.
- Aytekin, M., Grassi, E., Sahota, M., and Moss, C. F. (2004). "The bat head-related transfer function reveals binaural cues for sound localization in azimuth and elevation," *The Journal of the Acoustical Society of America* **116**, 3594-3605.
- Bajec, I. L., and Heppner, F. H. (2009). "Organized flight in birds," *Animal Behaviour* **78**, 777-789.

- Bates, M. E., Simmons, J. A., and Zorikov, T. V. (2011). "Bats use echo harmonic structure to distinguish their targets from background clutter," *Science* **333**, 627-630.
- Brinkløv, S., Jakobsen, L., Ratcliffe, J. M., Kalko, E. K., and Surlykke, A. (2011). "Echolocation call intensity and directionality in flying short-tailed fruit bats, *Carollia perspicillata* (Phyllostomidae) a," *The Journal of the Acoustical Society of America* **129**, 427-435.
- Britton, A., and Jones, G. (1999). "Echolocation behaviour and prey-capture success in foraging bats: laboratory and field experiments on *Myotis daubentonii*," *Journal of Experimental Biology* **202**, 1793-1801.
- Buckingham, G., Milne, J. L., Byrne, C. M., and Goodale, M. A. (2015). "The size-weight illusion induced through human echolocation," *Psychological Science* **26**, 237-242.
- Chiu, C., Reddy, P. V., Xian, W., Krishnaprasad, P. S., and Moss, C. F. (2010). "Effects of competitive prey capture on flight behavior and sonar beam pattern in paired big brown bats, *Eptesicus fuscus*," *Journal of Experimental Biology* **213**, 3348-3356.
- Cotzin, M., and Dallenbach, K. M. (1950). "" Facial Vision:" The Role of Pitch and Loudness in the Perception of Obstacles by the Blind," *The American Journal of Psychology*, 485-515.
- Crompton, B., Thomason, J. C., and McLachlan, A. (2003). "Mating in a viscous universe: the race is to the agile, not to the swift," *Proceedings of the Royal Society of London B: Biological Sciences* **270**, 1991-1995.
- Crundall, D., Underwood, G., and Chapman, P. (1999). "Driving experience and the functional field of view," *Perception* **28**, 1075-1087.
- DeLong, C. M., Au, W. W., Harley, H. E., Roitblat, H. L., and Pytka, L. (2007a). "Human listeners provide insights into echo features used by dolphins (*Tursiops truncatus*) to discriminate among objects," *Journal of Comparative Psychology* **121**, 306.
- DeLong, C. M., Au, W. W., and Stamper, S. A. (2007b). "Echo features used by human listeners to discriminate among objects that vary in material or wall thickness: implications for echolocating dolphins," *The Journal of the Acoustical Society of America* **121**, 605-617.

- Eckmeier, D., Geurten, B. R., Kress, D., Mertes, M., Kern, R., Egelhaaf, M., and Bischof, H.-J. (2008). "Gaze strategy in the free flying zebra finch (*Taeniopygia guttata*)," *PloS one* **3**, e3956.
- Falk, B., Williams, T., Aytikin, M., and Moss, C. F. (2011). "Adaptive behavior for texture discrimination by the free-flying big brown bat, *Eptesicus fuscus*," *Journal of Comparative Physiology A* **197**, 491-503.
- Fujioka, E., Aihara, I., Sumiya, M., Aihara, K., and Hiryu, S. (2016). "Echolocating bats use future-target information for optimal foraging," *Proceedings of the National Academy of Sciences* **113**, 4848-4852.
- Fujioka, E., Aihara, I., Watanabe, S., Sumiya, M., Hiryu, S., Simmons, J. A., Riquimaroux, H., and Watanabe, Y. (2014). "Rapid shifts of sonar attention by *Pipistrellus abramus* during natural hunting for multiple prey," *The Journal of the Acoustical Society of America* **136**, 3389-3400.
- Fujioka, E., Mantani, S., Hiryu, S., Riquimaroux, H., and Watanabe, Y. (2011). "Echolocation and flight strategy of Japanese house bats during natural foraging, revealed by a microphone array system," *The Journal of the Acoustical Society of America* **129**, 1081-1088.
- Ghose, K., Horiuchi, T. K., Krishnaprasad, P., and Moss, C. F. (2006). "Echolocating bats use a nearly time-optimal strategy to intercept prey," *PLoS Biology* **4**, e108.
- Ghose, K., and Moss, C. F. (2003). "The sonar beam pattern of a flying bat as it tracks tethered insects," *The Journal of the Acoustical Society of America* **114**, 1120-1131.
- Ghose, K., and Moss, C. F. (2006). "Steering by hearing: a bat's acoustic gaze is linked to its flight motor output by a delayed, adaptive linear law," *Journal of Neuroscience* **26**, 1704-1710.
- Ghose, K., Triplehorn, J. D., Bohn, K., Yager, D. D., and Moss, C. F. (2009). "Behavioral responses of big brown bats to dives by praying mantises," *Journal of Experimental Biology* **212**, 693-703.
- Gracheva, E. O., Ingolia, N. T., Kelly, Y. M., Cordero-Morales, J. F., Hollopeter, G., Chesler, A. T., Sánchez, E. E., Perez, J. C., Weissman, J. S., and Julius, D. (2010). "Molecular basis of infrared detection by snakes," *Nature* **464**, 1006-1011.

- Griffin, D. R. (1944). "Echolocation by blind men, bats and radar," *Science* **100**, 589-590.
- Hammer Jr, C. E., and Au, W. W. (1980). "Porpoise echo - recognition: An analysis of controlling target characteristics," *The Journal of the Acoustical Society of America* **68**, 1285-1293.
- Harley, H. E., Putman, E. A., and Roitblat, H. L. (2003). "Bottlenose dolphins perceive object features through echolocation," *Nature* **424**, 667-669.
- Hausfeld, S., Power, R. P., Gorta, A., and Harris, P. (1982). "Echo perception of shape and texture by sighted subjects," *Perceptual and Motor Skills* **55**, 623-632.
- Hemelrijk, C. K., and Kunz, H. (2004). "Density distribution and size sorting in fish schools: an individual-based model," *Behavioral Ecology* **16**, 178-187.
- Hildenbrandt, H., Carere, C., and Hemelrijk, C. K. (2010). "Self-organized aerial displays of thousands of starlings: a model," *Behavioral Ecology* **21**, 1349-1359.
- Hirai, T., and Kimura, S. (2004). "Diet composition of the common bat *Pipistrellus abramus* (Chiroptera; Vespertilionidae), revealed by fecal analysis," *Japanese Journal of Ecology (Japan)*.
- Hiryu, S., Hagino, T., Fujioka, E., Riquimaroux, H., and Watanabe, Y. (2008a). "Adaptive echolocation sounds of insectivorous bats, *Pipistrellus abramus*, during foraging flights in the field," *The Journal of the Acoustical Society of America* **124**, EL51-EL56.
- Hiryu, S., Hagino, T., Riquimaroux, H., and Watanabe, Y. (2007). "Echo-intensity compensation in echolocating bats (*Pipistrellus abramus*) during flight measured by a telemetry microphone," *The Journal of the Acoustical Society of America* **121**, 1749-1757.
- Hiryu, S., Katsura, K., Lin, L.-K., Riquimaroux, H., and Watanabe, Y. (2005). "Doppler-shift compensation in the Taiwanese leaf-nosed bat (*Hipposideros terasensis*) recorded with a telemetry microphone system during flight," *The Journal of the Acoustical Society of America* **118**, 3927-3933.
- Hiryu, S., Shiori, Y., Hosokawa, T., Riquimaroux, H., and Watanabe, Y. (2008b). "On-board telemetry of emitted sounds from free-flying bats: compensation for velocity

and distance stabilizes echo frequency and amplitude," *Journal of Comparative Physiology A* **194**, 841-851.

Hughes, B. (2001). "Active artificial echolocation and the nonvisual perception of aperture passability," *Human movement science* **20**, 371-400.

Ifukube, T., Sasaki, T., and Peng, C. (1991). "A blind mobility aid modeled after echolocation of bats," *IEEE Transactions on Biomedical Engineering* **38**, 461-465.

Jakobsen, L., Brinkløv, S., and Surlykke, A. (2013). "Intensity and directionality of bat echolocation signals," *Frontiers in physiology* **4**.

Jakobsen, L., Olsen, M. N., and Surlykke, A. (2015). "Dynamics of the echolocation beam during prey pursuit in aerial hawking bats," *Proceedings of the National Academy of Sciences* **112**, 8118-8123.

Jakobsen, L., and Surlykke, A. (2010). "Vespertilionid bats control the width of their biosonar sound beam dynamically during prey pursuit," *Proceedings of the National Academy of Sciences* **107**, 13930-13935.

Kalko, E. K., and Schnitzler, H.-U. (1989). "The echolocation and hunting behavior of Daubenton's bat, *Myotis daubentoni*," *Behavioral Ecology and Sociobiology* **24**, 225-238.

Kay, L. (1964). "An ultrasonic sensing probe as a mobility aid for the blind," *Ultrasonics* **2**, 53-59.

Kay, L. (1974). "A sonar aid to enhance spatial perception of the blind: engineering design and evaluation," *Radio and Electronic Engineer* **44**, 605-627.

Kay, L. (2000). "Auditory perception of objects by blind persons, using a bioacoustic high resolution air sonar," *The Journal of the Acoustical Society of America* **107**, 3266-3275.

Kellog, W. N. (1962). "Sonar system of the blind," *Science* **137**, 399-404.

Kick, S. A. (1982). "Target-detection by the echolocating bat, *Eptesicus fuscus*," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* **145**, 431-435.

- Kolarik, A. J., Cirstea, S., Pardhan, S., and Moore, B. C. (2014). "A summary of research investigating echolocation abilities of blind and sighted humans," *Hear Res* **310**, 60-68.
- Kounitsky, P., Rydell, J., Amichai, E., Boonman, A., Eitan, O., Weiss, A. J., and Yovel, Y. (2015). "Bats adjust their mouth gape to zoom their biosonar field of view," *Proceedings of the National Academy of Sciences* **112**, 6724-6729.
- Kuc, R. (1994). "Sensorimotor model of bat echolocation and prey capture," *The Journal of the Acoustical Society of America* **96**, 1965-1978.
- Kuc, R. (2012). "Echolocation with bat buzz emissions: Model and biomimetic sonar for elevation estimation," *The Journal of the Acoustical Society of America* **131**, 561-568.
- Land, M. F., and Tatler, B. W. (2001). "Steering with the head: The visual strategy of a racing driver," *Current Biology* **11**, 1215-1220.
- Liang, M., and Palakal, M. J. (1997). "A multiple target acoustic scene representation model for bat echolocation signals," (ASA).
- Matsuta, N., Hiryu, S., Fujioka, E., Yamada, Y., Riquimaroux, H., and Watanabe, Y. (2013). "Adaptive beam-width control of echolocation sounds by CF-FM bats, *Rhinolophus ferrumequinum nippon*, during prey-capture flight," *Journal of Experimental Biology* **216**, 1210-1218.
- Mihajlik, P., Guttermuth, M., Seres, K., and Tatai, P. (2001). "DSP-based ultrasonic navigation aid for the blind," *Instrumentation and Measurement Technology Conference, 2001. IMTC 2001. Proceedings of the 18th IEEE* **3**, 1535-1540.
- Milne, J. L., Anello, M., Goodale, M. A., and Thaler, L. (2015a). "A blind human expert echolocator shows size constancy for objects perceived by echoes," *Neurocase* **21**, 465-470.
- Milne, J. L., Arnott, S. R., Kish, D., Goodale, M. A., and Thaler, L. (2015b). "Parahippocampal cortex is involved in material processing via echoes in blind echolocation experts," *Vision Research* **109**, 139-148.
- Milne, J. L., Goodale, M. A., and Thaler, L. (2014). "The role of head movements in the

discrimination of 2-D shape by blind echolocation experts," *Atten Percept Psychophys* **76**, 1828-1837.

Moriya, N., and Kaneda, Y. (2005). "Study of harmonic distortion on impulse response measurement with logarithmic time stretched pulse," *Acoustical science and technology* **26**, 462-464.

Motoi, K., Sumiya, M., Fujioka, E., and Hiryu, S. (2017). "Three-dimensional sonar beam-width expansion by Japanese house bats (*Pipistrellus abramus*) during natural foraging," *The Journal of the Acoustical Society of America* **141**, EL439-EL444.

Pelegrín-García, D., Rychtáriková, M., and Glorieux, C. (2015). "Human echolocation: localizing reflections of self-generated oral sounds in laboratory."

Rice, C. E. (1967). "Human echo perception," *Science* **155**, 656-664.

Rice, C. E. (1969). "Perceptual enhancement in the early blind?," *The Psychological Record* **19**, 1.

Rice, C. E., and Feinstein, S. H. (1965). "Sonar system of the blind: size discrimination," *Science* **148**, 1107-1108.

Rojas, J. A. M., Hermosilla, J. A., Montero, R. S., and Espí, P. L. L. (2009). "Physical analysis of several organic signals for human echolocation: oral vacuum pulses," *Acta acustica united with acustica* **95**, 325-330.

Rojas, J. A. M., Hermosilla, J. A., Montero, R. S., and Espí, P. L. L. (2010). "Physical analysis of several organic signals for human echolocation: hand and finger produced pulses," *Acta acustica united with acustica* **96**, 1069-1077.

Rowan, D., Papadopoulos, T., Edwards, D., and Allen, R. (2015). "Use of binaural and monaural cues to identify the lateral position of a virtual object using echoes," *Hear Res* **323**, 32-39.

Rowan, D., Papadopoulos, T., Edwards, D., Holmes, H., Hollingdale, A., Evans, L., and Allen, R. (2013). "Identification of the lateral position of a virtual object based on echoes by humans," *Hear Res* **300**, 56-65.

Ryu, D., Abernethy, B., Mann, D. L., and Poolton, J. M. (2015). "The contributions of central and peripheral vision to expertise in basketball: How blur helps to provide

a clearer picture," *Journal of Experimental Psychology: Human Perception and Performance* **41**, 167.

Schall, R. (1991). "Estimation in generalized linear models with random effects," *Biometrika* **78**, 719-727.

Schaub, A., and Schnitzler, H.-U. (2007). "Flight and echolocation behaviour of three vespertilionid bat species while commuting on flyways," *Journal of Comparative Physiology A* **193**, 1185-1194.

Schenkman, B. N. (1986). "Identification of ground materials with the aid of tapping sounds and vibrations of long canes for the blind," *Ergonomics* **29**, 985-998.

Schenkman, B. N., and Nilsson, M. E. (2010). "Human echolocation: Blind and sighted persons' ability to detect sounds recorded in the presence of a reflecting object," *Perception* **39**, 483-501.

Schenkman, B. N., and Nilsson, M. E. (2011). "Human echolocation: pitch versus loudness information," *Perception* **40**, 840-852.

Schenkman, B. N., Nilsson, M. E., and Grbic, N. (2016). "Human echolocation: Acoustic gaze for burst trains and continuous noise," *Applied Acoustics* **106**, 77-86.

Schnitzler, H.-U., Kalko, E., Miller, L., and Surlykke, A. (1987). "The echolocation and hunting behavior of the bat, *Pipistrellus kuhli*," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* **161**, 267-274.

Schnitzler, H.-U., and Kalko, E. K. (2001). "Echolocation by Insect-Eating Bats: We define four distinct functional groups of bats and find differences in signal structure that correlate with the typical echolocation tasks faced by each group," *Bioscience* **51**, 557-569.

Schornich, S., Nagy, A., and Wiegrebe, L. (2012). "Discovering your inner bat: echo-acoustic target ranging in humans," *J Assoc Res Otolaryngol* **13**, 673-682.

Seibert, A.-M., Koblitz, J. C., Denzinger, A., and Schnitzler, H.-U. (2013). "Scanning behavior in echolocating common pipistrelle bats (*Pipistrellus pipistrellus*)," *PloS one* **8**, e60752.

Simmons, J., Kick, S., Lawrence, B., Hale, C., Bard, C., and Escudie, B. (1983). "Acuity

of horizontal angle discrimination by the echolocating bat, *Eptesicus fuscus*," *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology* **153**, 321-330.

Simmons, J., Lavender, W., Lavender, B., Doroshov, C., Kiefer, S., Livingston, R., Scallet, A., and Crowley, D. (1974). "Target structure and echo spectral discrimination by echolocating bats," *Science* **186**, 1130-1132.

Simmons, J. A. (1973). "The resolution of target range by echolocating bats," *The Journal of the Acoustical Society of America* **54**, 157-173.

Simmons, J. A. (1979). "Perception of echo phase information in bat sonar," *Science* **204**, 1336-1338.

Simmons, J. A. (2005). "Big brown bats and June beetles: multiple pursuit strategies in a seasonal acoustic predator-prey system," *Acoustics Research Letters Online* **6**, 238-242.

Simmons, J. A., Fenton, M. B., and O'Farrell, M. J. (1979). "Echolocation and pursuit of prey by bats," *Science* **203**, 16-21.

Simmons, N. B., Seymour, K. L., Habersetzer, J., and Gunnell, G. F. (2008). "Primitive Early Eocene bat from Wyoming and the evolution of flight and echolocation," *Nature* **451**, 818-821.

Sohl-Dickstein, J., Teng, S., Gaub, B. M., Rodgers, C. C., Li, C., DeWeese, M. R., and Harper, N. S. (2015). "A Device for Human Ultrasonic Echolocation," *IEEE Transactions on Biomedical Engineering* **62**, 1526-1534.

Stan, G.-B., Embrechts, J.-J., and Archambeau, D. (2002). "Comparison of different impulse response measurement techniques," *Journal of the Audio Engineering Society* **50**, 249-262.

Sumiya, M., Fujioka, E., Motoi, K., Kondo, M., and Hiryu, S. (2017). "Coordinated Control of Acoustical Field of View and Flight in Three-Dimensional Space for Consecutive Capture by Echolocating Bats during Natural Foraging," *PloS one* **12**, e0169995.

Sumiya, M., Fujioka, E., Watanabe, Y., Hiryu, S., and Aihara, I. (2015). "Mathematical

modeling of flight and acoustic dynamics of an echolocating bat during multiple-prey pursuit," Proceedings of 2015 International Symposium on Nonlinear Theory and its Applications (NOLTA2015), pp.369–372.

Supa, M., Cotzin, M., and Dallenbach, K. M. (1944). " " Facial vision": the perception of obstacles by the blind," The American Journal of Psychology **57**, 133-183.

Surlykke, A., Ghose, K., and Moss, C. F. (2009a). "Acoustic scanning of natural scenes by echolocation in the big brown bat, *Eptesicus fuscus*," Journal of Experimental Biology **212**, 1011-1020.

Surlykke, A., and Moss, C. F. (2000). "Echolocation behavior of big brown bats, *Eptesicus fuscus*, in the field and the laboratory," The Journal of the Acoustical Society of America **108**, 2419-2429.

Surlykke, A., Pedersen, S. B., and Jakobsen, L. (2009b). "Echolocating bats emit a highly directional sonar sound beam in the field," Proceedings of the Royal Society of London B: Biological Sciences **276**, 853-860.

Teng, S., Puri, A., and Whitney, D. (2012). "Ultrafine spatial acuity of blind expert human echolocators," Experimental Brain Research **216**, 483-488.

Teng, S., and Whitney, D. (2011). "The acuity of echolocation: Spatial resolution in the sighted compared to expert performance," J Vis Impair Blind **105**, 20-32.

Thaler, L. (2015). "Using sound to get around - discoveries in human echolocation," Observer **28**.

Thaler, L., Arnott, S. R., and Goodale, M. A. (2011). "Neural correlates of natural human echolocation in early and late blind echolocation experts," PloS one **6**, e20162.

Thaler, L., and Castillo-Serrano, J. (2016). "People's Ability to Detect Objects Using Click-Based Echolocation: A Direct Comparison between Mouth-Clicks and Clicks Made by a Loudspeaker," PloS one **11**, e0154868.

Thaler, L., and Goodale, M. A. (2016). "Echolocation in humans: an overview," Wiley Interdiscip Rev Cogn Sci **7**, 382-393.

Thaler, L., Reich, G. M., Zhang, X., Wang, D., Smith, G. E., Tao, Z., Abdullah, R. S. A. B. R., Cherniakov, M., Baker, C. J., and Kish, D. (2017). "Mouth-clicks used by

blind expert human echolocators—signal description and model based signal synthesis," *PLoS Computational Biology* **13**, e1005670.

Tian, B., and Schnitzler, H.-U. (1997). "Echolocation signals of the greater horseshoe bat (*Rhinolophus ferrumequinum*) in transfer flight and during landing," *The Journal of the Acoustical Society of America* **101**, 2347-2364.

Uchibori, S., Sarumaru, Y., Ashihara, K., Ohta, T., and Hiryu, S. (2015). "Experimental evaluation of binaural recording system using a miniature dummy head," *Acoustical Science and Technology* **36**, 42-45.

Vicsek, T., and Zafeiris, A. (2012). "Collective motion," *Physics Reports* **517**, 71-140.

von der Emde, G. (2006). "Non-visual environmental imaging and object detection through active electrolocation in weakly electric fish," *Journal of Comparative Physiology A* **192**, 601-612.

von der Emde, G., and Schnitzler, H.-U. (1990). "Classification of insects by echolocating greater horseshoe bats," *Journal of Comparative Physiology A* **167**, 423-430.

Wallmeier, L., Gessele, N., and Wiegrebe, L. (2013). "Echolocation versus echo suppression in humans," *Proc Biol Sci* **280**, 20131428.

Wallmeier, L., and Wiegrebe, L. (2014a). "Ranging in human sonar: Effects of additional early reflections and exploratory head movements," *PloS one* **9**, e115363.

Wallmeier, L., and Wiegrebe, L. (2014b). "Self-motion facilitates echo-acoustic orientation in humans," *Royal Society open science* **1**, 140185.

Wallmeier, L., and Wiegrebe, L. (2014c). "Self-motion facilitates echo-acoustic orientation in humans," *R Soc Open Sci* **1**, 140185.

Waters, D. A., and Abulula, H. H. (2007). "Using bat-modelled sonar as a navigational tool in virtual environments," *International Journal of Human-Computer Studies* **65**, 873-886.

Williams, A., and Davids, K. (1998). "Visual search strategy, selective attention, and expertise in soccer," *Research quarterly for exercise and sport* **69**, 111-128.

Worchel, P., and Dallenbach, K. M. (1947). "" Facial Vision:" Perception of Obstacles

by the Deaf-Blind," *The American Journal of Psychology* **60**, 502-553.

Wotton, J. M., Jenison, R. L., and Hartley, D. J. (1997). "The combination of echolocation emission and ear reception enhances directional spectral cues of the big brown bat, *Eptesicus fuscus*," *The Journal of the Acoustical Society of America* **101**, 1723-1733.

Yovel, Y., Falk, B., Moss, C. F., and Ulanovsky, N. (2011). "Active control of acoustic field-of-view in a biosonar system," *PLoS Biology* **9**, e1001150.

Yovel, Y., Franz, M. O., Stilz, P., and Schnitzler, H.-U. (2008). "Plant classification from bat-like echolocation signals," *PLoS Comput Biol* **4**, e1000032.

Curriculum Vitae

Miwa Sumiya

Date of birth

August 29th, 1990

Place of birth

Kyoto, Japan

Education

Doshisha University in Kyotanabe:

B.Sc. Engineering, March 2013

M.Sc. Engineering, March 2015

Professional Experience

April 2015–March 2018. Research Fellow of Japan Society for the Promotion of Science (DC1) in Doshisha University

Professional members

The Acoustical Society of Japan

The Acoustical Society of America

November 2017