UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

Institute of Sound and Vibration Research

Development and Application of an Echolocation Model Inspired by Bats

by

Suyeon Kim

thesis for the degree of Doctor of Philosophy

May, 2010

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

INSTITUTE OF SOUND AND VIBRATION RESEARCH

Doctor of Philosophy

DEVELOPMENT AND APPLICATION OF AN ECHOLOCATION MODEL INSPIRED BY BATS

by Suyeon Kim

This study explores the principles of echolocation in bats which can be potentially adopted for bio-inspired sonar systems. Using a biological signal processing technique which was developed based on bat's hearing system, the effect of auditory processing on the object discrimination is investigated for both CF (constant frequency) and FM (frequency modulated) signals respectively. These signals are considered as two representative types of echolocating calls. This study has simulated returning echoes from target discs using different types of calls by applying measured impulse responses of the objects. The simulated echoes were then processed through auditory models. The results have shown that the auditory processing contributes not only to increase the gain but also to enhance the ability to discriminate the sizes of discs. The peak and notch characteristics appearing in the auditory spectrum also confirms the flexibility of designing auditory models to manipulate spectral and temporal characteristics of the echo signals. Secondly, the effect of the bat's head on the received signals at the two ears for varying distances was investigated by measuring the head-related transfer function (HRTF) of a bat-head cast. It has been reported that a bat changes bandwidth and duration of its echolocating call as it approaches a target. Adaptive change in the echolocating calls has been well explained in previous studies in terms of characteristics of signal structure. However, the range-dependent adaptive change in emitted signals also implies that the reflected signals reaching the two ears (i.e. binaural hearing) change in gain and frequency as the distance between the bat and the target varies. The result of measured HRTF has provided insights to range-dependent binaural information regarding the adaptive change of the echolocating calls. The results of measured

data show that relatively higher gain at low frequencies (below 10 kHz[1]) is observed than that at high frequencies (above 10 kHz) as the bat-head cast approaches the sound source. It is also noted that interaural level differences (ILDs) at a fixed distance have less sensitive changes at low frequencies than at high frequencies as the angle of the source direction changes in the frontal axis. However, the sensitivity of the ILDs at low frequencies increase more than at high frequencies as the range reduces. It is concluded that the low frequency implies a more significant role during the target approaching stage in echolocation including distance perception. Also, the systematic change in sensitivity of the ILDs in various ranges suggests that the bat might be able to calibrate the angular resolution by broadening the bandwidth at low frequencies. Furthermore, the HRTF results calculated from a computational sphere model confirms the potential function of low frequency to calibrate the ILDs sensitivity for varying distances. Overall, this study has shown that customised auditory processing of the echolocating signal improves the quality of sonar representation and the results of investigations using the HRTFs of the bat-head cast guide the future design of effective adaptive signals based on the range-dependent HRTFs, to potentially enhance the performance of sonar systems.

[1]This study has defined the range of the low and the high frequencies based on the acoustical diffraction and reflection of the sound around the bat-head. The diffraction effect appeared to be prominent below 10 kHz.

Contents

List of Figures

List of Tables

Declaration of authorship

I, Suyeon Kim, declare that the thesis entitled "*Development and Application of an Echolocation Model Inspired by Bats*" and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

Signed: .....

Date: .....

Acknowledgements

First of all, I would like to acknowledge and greatly thank my PhD supervisor, Bob Allen. His continuous support, encouragement and academic advice have led me to complete the PhD. I would also like to thank my second supervisor, Daniel Rowan for his inspiring enthusiasm and invaluable help on comments and suggestions of the dissertation. For financial support, I am grateful to Research Council UK which is the sponsor of the project, and ISVR department which has offered me the Rayleigh scholarship. I also thank Dave, Timos, Munhum and Dragana for advice, opinions and encouragement while we were working on BIAS (Biologically Inspired Acoustic Systems) project together. I would also like to acknowledge Alicia who helped experiments for days in the anechoic chamber. I am also grateful to people in ISVR workshop and mechanical support team who helped me build the measurement rig and make the bat-head cast.

I would like to thank Kitchener family for invaluable friendship and film nights. I thank Korean church in Southampton for support and prayers. I thank Nick for your love, support and most delicious lunch boxes. Finally, I would like to dedicate this thesis to my family. Thanks Mum, Dad and my little sister Suji for your limitless support and encouragement, and for parcels and e-mails from miles away.

Acknowledgements of Sponsors

The author gratefully acknowledges the support of Research Councils UK (RCUK) through its Basic Technology Programme and the contributions of the consortium partners from Universities of Edinburgh, Leeds and Leicester, and from British Geological Survey and from Fortkey Ltd.

Abbreviations

CF	Constant frequency
ERB	Equivalent rectangular bandwidth
FFT	Fast fourier transform
FM	Freqeuncy modulated
HRIR	Head-related impulse response
HRTF	Head-related transfer function
ILD	Interaural level difference
ITD	Interaural time difference
LP	Large plastic
LTI	Liner time invariant
MP	Medium plastic
OHC	Outer hair cells
PSD	Power spectral density
SP	Small plastic
SPL	Sound pressure level
TW	Thick wooden
tW	Thin wooden

Symbols

A	Amplitude of interaural HRTF in frequency domain
A_L	Amplitude of HRTF at left ear in frequency domain
A_R	Amplitude of HRTF at right ear in frequency domain
$EarQ$	Asymptotic filter quality
H	Interaural HRTF in frequency domain
H_L	HRTF at left ear in frequency domain
H_R	HRTF at right ear in frequency domain
M_L	Average gain at left ear
M_R	Average gain at right ear
N	Sample's number at discrete time notation
P_m	Legendre polynomial function
Q_{10}	Centre frequency divided by bandwidth 10 dB from the tip of the bandpass filter
S_ω	Source strength magnitude of flow from an ideal point source
a	Radius of head
a_g	Amplitude coefficient for gammatone impulse response
b	Bandwidth of filter
c	Speed of sound
d	Distance between the sound source and interaural centre
f	Frequency in Hz
f_c	Centre frequency in Hz
$g(t)$	Gammatone impulse response
h_L	HRTF at left ear in time domain
h_R	HRTF at right ear in time domain
$h_{bacscatter}$	Backscattering response
h_m	mth-order spherical Hankel function
k	Wave number

m	Degree of order in function
$minBW$	Minimum bandwidth
n	Filter order
p_{ear}	Signal reaching the ear drum
p_{echo}	Returning echolocating signal
p_{ff}	Free-field sound pressure
p_{out}	Emitted signal
p_s	Sound pressure on the surface of a sphere
$\mathbf{r}$	Position in vector
r	Distance
t	Time sequence
Φ	Interaural phase of HRTF
Φ_L	Phase of HRTF at left ear
Φ_R	Phase of HRTF at right ear
ρ	Normalised distance
θ	General notation for azimuth
θ_p	Perceived angle
φ	General notation for elevation
ϕ_g	Phase delay in gammatone impulse response
$\phi(\omega)$	Phase delay in time sequence
ρ_0	Density of air
σ_L	Standard deviation of average gain at left ear
σ_R	Standard deviation of average gain at right ear
τ	General notation for ITD
τ_G	Group delay
τ_ϕ	Phase delay in radian
$\tau_{\phi,theoretical}$	Theoretical value of phase delay
ω	Angular frequency

Chapter 1

Introduction

Echolocation is an acoustical process which is used to locate and identify a target by sending sound pulses and receiving the echoes reflected back from the target. Echolocation is used by a number of mammalian groups but the echolocation of bats has attracted the most scientific research. Most bats fly in complete darkness and use echolocation to capture their prey and to communicate between themselves by using self-generated biosonar signals. Since bat echolocation was first observed [Griffin, 1958], their ability has attracted much interest from researchers in various academic areas including biology, physiology and acoustics.

It has also attracted interest from engineers, who have been trying to learn from biological echolocating behaviour, and apply the principles to sonar engineering. Müller and Kuc [2007] have recently reviewed biosonar-inspired technology motivated by echolocating animals. In their review, the importance of insights into various echolocating tasks solved by bats has been stressed as it potentially provides a wide range of methodologies for engineering of any technical sonar system. For example, binaural tracking sensors and rotational sensors have been motivated by the hearing mechanism of two ears and movement of external ears in bats [Barshan and Kuc, 1992; Kuc, 1996; Carmena and Hallam, 2001; Carmena et al., 2001]. Also the signal processing to classify a target by scanning the echoes over the elevation is motivated by dolphins' behaviour of nodding up-and-down in echolocation [Kuc, 1997b]. Although there have been some studies about the implementation of the bat's sonar system in robotics, studies to develop an echolocating-task model based on bat acoustics or bat hearing is relatively scarce. The task modelling is considered to be important because it can be used to evaluate and improve the performance of sonar system by changing input variables of the modelling, however there are not many studies which model the task in the echolocation processing, computationally or experimentally. The echolocating-task modelling can be categorised into detection, discrimination and localisation as these are the tasks which

a bat is required to perform to catch a prey. Many studies in biology and physiology which reveal the behaviour of bats in solving these tasks inspires echolocating model.

In other words, the scientific research area which transfers the insight found in bat's biology to biosonar engineering has high potentials to be developed. In particular, an echolocating model is required to recreate the principles of echolocating performance in bats but not to blindly copy their physical characteristics. Therefore, the modelling technique is considered to be important as it contributes to the understanding of the bat's echolocating behaviour through engineering tools. It also helps to understand general principles of echolocation which can be adopted in sonar system design for the future. Also developing the idea of bio-inspired sonar in the engineering environment is considered necessary because the materials used in engineering are substantially different from those used in nature. Forbes [2006] recently expressed and important principle of bio-inspired engineering in this regard:

"Scientists try to unravel nature's mechanisms, but technologists use whatever will work. Bio-inspired technical products will almost certainly not mimic the actual materials used by nature."

Forbes emphasises the importance of finding global principles which can be adopted to design a system using technical methods. The technical methods may include mechanical/experimental set-ups, receiver design, and signal processing techniques, etc. In this study, the types of echolocating signals, auditory processing and binaural hearing are examined through either modelling or experiments. The broad aims of this study can be stated as follows:

(1) Develop a methodology to model the echolocating process (Chapter 2).
(2) Evaluate the performance of different types of echolocating signal on a discrimination task (Chapter 3).
(3) Investigate the potential benefits of adopting a bat's auditory model for the analysis of echolocating signals (Chapter 3).
(4) Develop experimental methodology to build a binaural processing platform (Chapter 4).
(5) Gain insights into the role of binaural hearing for 'active' localisation (echolocation) (Chapter 5).

Also, the overall structure of this study is described in Figure 1.1.

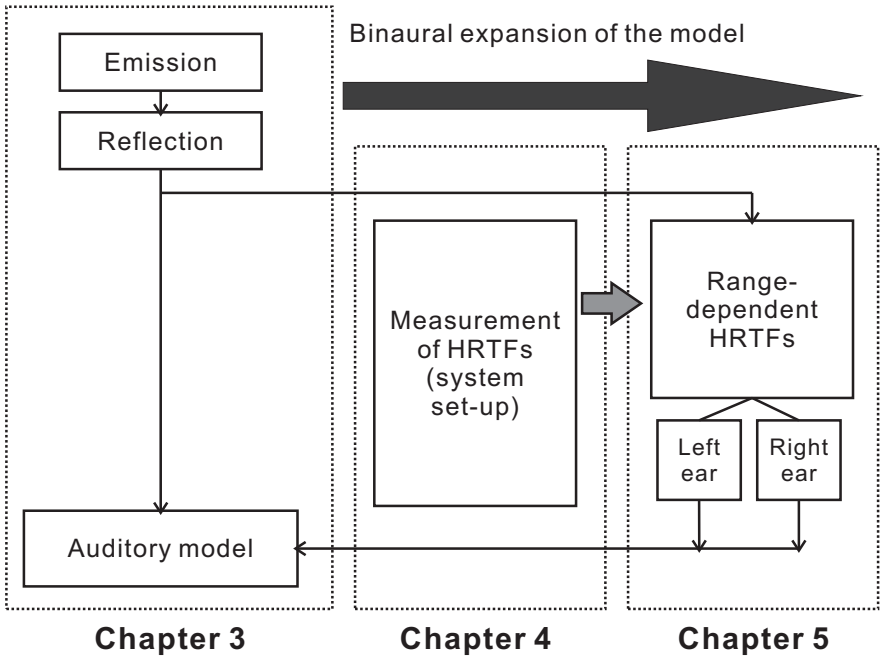


Figure 1.1: Overall structure of this study. The modelling of echolocation and its application is described in Chapter 3, followed by the development of a head-related transfer function (HRTF) of a bat-head cast for binaural studies in Chapter 4. The acoustic characteristics of the bat-head HRTF are investigated at a range of distances in Chapter 5.

To model echolocation, it is useful to separate each stage of the process. Emission, reflection and reception are the basic components which comprise the model. In the emission stage, there are several characteristics of echolocating call such as signal type, duration and directivity, which may vary depending on bat species. On the other hand, the type of foliage where the sound is transmitted, the type of target and reflecting pattern, the reduction in signal amplitude in echoes and interference from signals of other bats are considered to be factors in designing the reflection stage. The shape and feature of the external ears and movements of pinna are important factors in designing the reception stage. Recently, a method to model the echolocation processing has been suggested to simulate the echoes returning back from a specific target in Institute of Sound and Vibration Research (ISVR) [Papadopoulos and Allen, 2007]. The concept is that each echolocating stage of the signal emission, reflection and reception is considered to be an individual modelling 'box'. Each stage is modelled as a linear and time invariant system. These stages are then combined in series to simulate the overall echolocation processing. The modelled 'boxes' can be combined by using a mathematical technique called convolution. Convolution is a mathematical operation on two different inputs to produce a third output which contains the effects of two linear func-

tions. Using this operation, echoes which bats might receive from a specific target can be predicted. This modelling approach enables us to concentrate on individual components of echolocation, and thus allows us to isolate the effects of changes in individual stages.

In Chapter 3, a bat emission is modelled by using two different types of echolocating signals, constant frequency (CF) and frequency modulated (FM), and the signals are digitally synthesised. The reflection component is obtained by measuring backscattering impulse responses from different objects. The backscattering impulse response is obtained by recording the reflected signal from a target when a broadband signal is sent towards the target. It contains information on the target in both time and frequency domains. Discs of various sizes and materials were used to obtain the responses. The receptive signal can be predicted by convolving the pre-designed emitted signal and pre-obtained response of the target. The reception is modelled as a single receptor and it is assumed to be in a central position in front of the bat's head.

Further modelling work was carried out in terms of auditory processing. The modelling is primarily based on the mammalian auditory system but a modification in the auditory filter design is introduced to include specific features for the bat (see below). The input to the auditory processing mostly refers to the signal which travels around the head and ears, then reaches the eardrum and the inner ear. However, the modelling environment is restricted to single receiver, hence the the predicted signal at the single receptor is directly applied to the auditory processing model. Note that the binaural reception is investigated separately in the following Chapter 4 and Chapter 5.

As mentioned earlier, it is important to develop a method to understand the principles of echolocation in bats, and the modelling work described above provides the platform to examine the principles. The first gap in knowledge is that the performance of artificial echolocating model to discriminate real experimental targets has not shown. Most research up to date has been behavioural studies. Therefore, previous studies have been dependent on a specific type of bat and its diet. The investigation in the artificial model is expected to provide result which is independent from biological and behavioural features of a living bat. Secondly, studies which examine echolocating performance of a particular task using a computational or non-behavioural method are scarce. Evaluating the echolocating performance in the artificial model is considered effective for implementation in potential biosonar applications. Therefore, this study aims to utilise the proposed model and examine its performance on a given object

discrimination task. This study has examined the performance of different types of echolocating signals, and auditory processing in echolocation.

The target discrimination task was applied in the suggested model and the results are shown in Chapter 3. Using the CF and FM signals, echoes received at the single receptor for the different experimental targets were predicted and compared with each other. The performance of discrimination observed in the results was investigated for each type of signal. Next a model of auditory processing was applied. The auditory processing model is separately designed for the CF and FM signals. For FM signals, the model is based on the general mammalian auditory system but the degree of auditory performance in resolving different frequencies is varied and the results are discussed. CF bats, however, are known to have a specialised hearing processes at the inner ear, which resolves the frequencies very finely around the echolocating frequency. Hence, the hearing model of CF bats is designed to possess a narrow and finely tuned frequency analysis region (called an *acoustic fovea*) around the given echolocating frequency.

One contribution of this study is the implementation of the object discrimination task within the echolocating model. The results show that different objects can be discriminated by the various types of artificial bat's signals being generated in the simulation although their performances vary depending on the types of signals. It also shows that the performance of the different types of signals varied according to their spectral characteristics. It is already known from previous studies that the CF signal is utilised for Doppler- shift compensation whereas the FM signal suits spectral analysis of target reflections in bat echolocation. Using the results shown in this study, potential advantages and disadvantages of CF and FM signals are proposed, which agree with previous studies based on alternative analysis. Therefore, object discrimination in the echolocating process is able to be recreated and assessed in terms of the function of different echolocating signals within the single receptor model.

However, it is necessary to model two receptors which represent the two ears in the bat in order to understand binaural processing, as sound localisation is mainly dominated by interaural (between-the-ears) differences. Ideally, a head-related impulse response (HRIR) and a head-related transfer function (HRTF) contains the time and frequency information of the signal reaching each ear respectively, and the difference between the two ears provides information of binaural hearing. It is conventionally obtained by the measurement of the responses from the head, external ear, and/or body at different azimuths and elevations at a specified distance.

In Chapter 4, this study introduces three major issues in relation to the HRTF of bats to provide better understanding of designing binaural receivers for sonar systems. They are listed as practical techniques in HRTF measurement, general characteristics of HRTFs and range-dependent characteristics of HRTFs. Firstly, there have been a few studies on the measurement of HRTFs in real bats [Obrist et al., 1993; Jen and Chen, 1988; Fuzessery, 1996; Aytekin et al., 2004; Firzlaff and Schuller, 2004]. The hardware set-up and the methodology of the HRTF measurement are considered to be important as they should cover the bat sonar frequency range which often reaches around or over 100 kHz. With small bat head sizes, it is difficult to measure this response. The measured data is expected to contain non-negligible errors if the hardware set-up is not reliable. Therefore, the previous studies of the methodology of the HRTFs measurement in small-headed animals were extensively reviewed and potential measurement errors were discussed. Hence, the system used in this thesis was developed based on the knowledge gained from the literature review and the gaps proposed in this study. Also the measurement system was examined to verify its performance. Then the HRTFs of an artificial bat-head cast were measured at a fixed distance from the sound source and the repeatability is examined. Secondly, the characteristics in the measured HRTFs of a bat-head cast in different azimuths are described. This is intended to extract the general features which are based on the physical size of the head and the frequency range of the system used. The acoustics of the sound travelling around the artificial head and pinna are discussed. Eventually it is aimed to develop a robust platform which enables us to measure the HRTFs and develop the binaural receptors. Finally, in Chapter 5, the range-dependent characteristics of HRTFs of the bat-head cast are investigated to provide better insight into the potential effect of binaural processing in an active sensing echolocating system. Considering the fact that echolocation is based on active movement during a bat's flight, the HRTFs measured at fixed position does not reflect the realistic environment of binaural receivers. By measuring the HRTFs of a bat-head cast for the sound sources at different distances, a moving sensor which approaches a target source is recreated. It contributes to understanding why binaural receivers are beneficial for echolocation when the receivers are active in motion.

The contents of this thesis is summarised as follows. In Chapter 2, a literature review will be presented in three major categories: overview of the bat, characteristics of the echolocating signals and the behaviour performance in the echolocating task observed in bats. Then the concept of single and binaural receptors is briefly introduced. In Chapter 3, the modelling work on the single receptor and the simulation of object

discrimination is presented and the results are discussed. In Chapter 4, the work on developing the method for the measurement of HRTFs in the bat is presented and the results of measuring the HRTFs of a bat-head cast are discussed. In Chapter 5, the characteristics of HRTFs measured at various distances are presented. The sphere model prediction was carried out to validate the result and compare the measured data to the computational data. In Chapter 6, a general discussion is made to provide possible issues in designing bat-inspired sonar systems based on the findings of this research. Finally, the conclusions and the future work are presented in Chapter 7.

The contribution to knowledge made by this research is as follows.

(1) This study has discovered that the convolution model of the echolocation can be used to develop a tool to assess an echolocating task. It was able to evaluate the object discrimination performance using this model.
(2) It has been discovered that the auditory processing of the echolocating signals improves the performance of object discrimination in terms of both gain and target characterisation.
(3) The result of measurement of a bat-head cast has revealed that type of microphone set-up and accuracy of interaural axis adjustment affects the measured data significantly, particularly at high frequencies above 60 - 70 kHz.
(4) This study has not only confirmed that the binaural characteristics of the bat-head cast are general across other kinds of mammals with two ears, but also the characteristics depends on head-size, frequency and distance between the source and the receiver.
(5) The measured HRTFs have demonstrated how the reduction in distance affects monaural/binaural information, and shown the dependency of interaural level differences (ILDs) on both frequency and distance which suggests a role of adaptive ILDs to calibrate the angular sensitivity.
(6) The thesis also provides both experimental and computational evidence to support the adaptive change in frequency of echolocating calls as a bat approaches its target.
(7) The measured inteaural time difference (ITDs), however, were not shown to be reliable for the lateral sound sources when distance varies, but it was shown that the ITDs can be potentially used to predict the angular position of a sound source for azimuth within $\pm 30^\circ$ of the midsagittal plane.
(8) The calibration effects derived from binaural information contribute to providing potential guidelines for frequency, size of head and distance when one designs binaural sonar receivers.

Parts of this study have been reported at following international conferences.

- S. Y. Kim, T. Papadopoulos, R. Allen, D. Rowan,"Using a model pf bat echolocation to differentiate object dimensions and material", Proceedings of the Biological Approaches for Engineering Conference, 17-19 March 2008
- S. Y. Kim, R Allen and D. Rowan, "The simulation of bat-oriented auditory processing using the experimental data of echolocating signal", Acoustics' 08, 29 June-4 July 2008, Paris, France
- S. Y. Kim, R Allen and D. Rowan,"Modelling of bat echolocation signals to differentiate object dimensions and material", 15th International Congress on Sound and Vibration, 6-10 July 2008, Daejeon, Korea
- S. Y. Kim, T. Papadopoulos, R. Allen and D. Rowan, "Measurement of the head-related transfer function of a bat-head cast", Proceedings of 5th International Conference on Bio-Acoustics, Loughborough, 2009, pp. 167-174

Chapter 2

Background

2.1 Introduction

Many studies have investigated echolocating mechanisms in bats. These include behavioural studies in experimental biology, anatomical studies in neurophysiology, simulations of computational models in echolocation and even mechanical modelling of robotic sensorimotor systems inspired by bats. Extensive reviews support the literature. Some of the key papers are summarised here with the intention of helping the reader to understand the echolocation processing model used within this study and discussed later in this chapter. As shown in Figure 2.1, the echolocating process can be viewed as a combination of the characteristics of the bat, the signal transmission, and the tasks involved in capturing target prey. Therefore, literature on echolocation is divided into three different categories; biological backgrounds of bats (A), echolocation and signals produced (B), and the echolocating tasks (C). In section 2.2, the biological background is reviewed in terms of species, history of evolution, and more importantly, anatomy and hearing perception which include the anatomical evidence of the echolocating skills being observed in the hearing system of the bat. In section 2.3, the definition of echolocation is discussed and the different types of signals are described. In section 2.4, the three major tasks of detection, discrimination and localisation, which are required to achieve successful echolocation, are discussed along with the abilities shown from previous behavioural studies.

Finally, two echolocation processing models are introduced and described in section 2.5 building on the earlier stages. Figure 2.1 shows the simple concept of a single receptor model (D) and a binaural receptor model (E). The basic components of the model are a signal generator, outgoing signal transmission (from the bat to the target), target reflection, echo transmission (from the target to the bat) and signal reception. The signal reception part is more complicated as there exist effects from the bat's head and

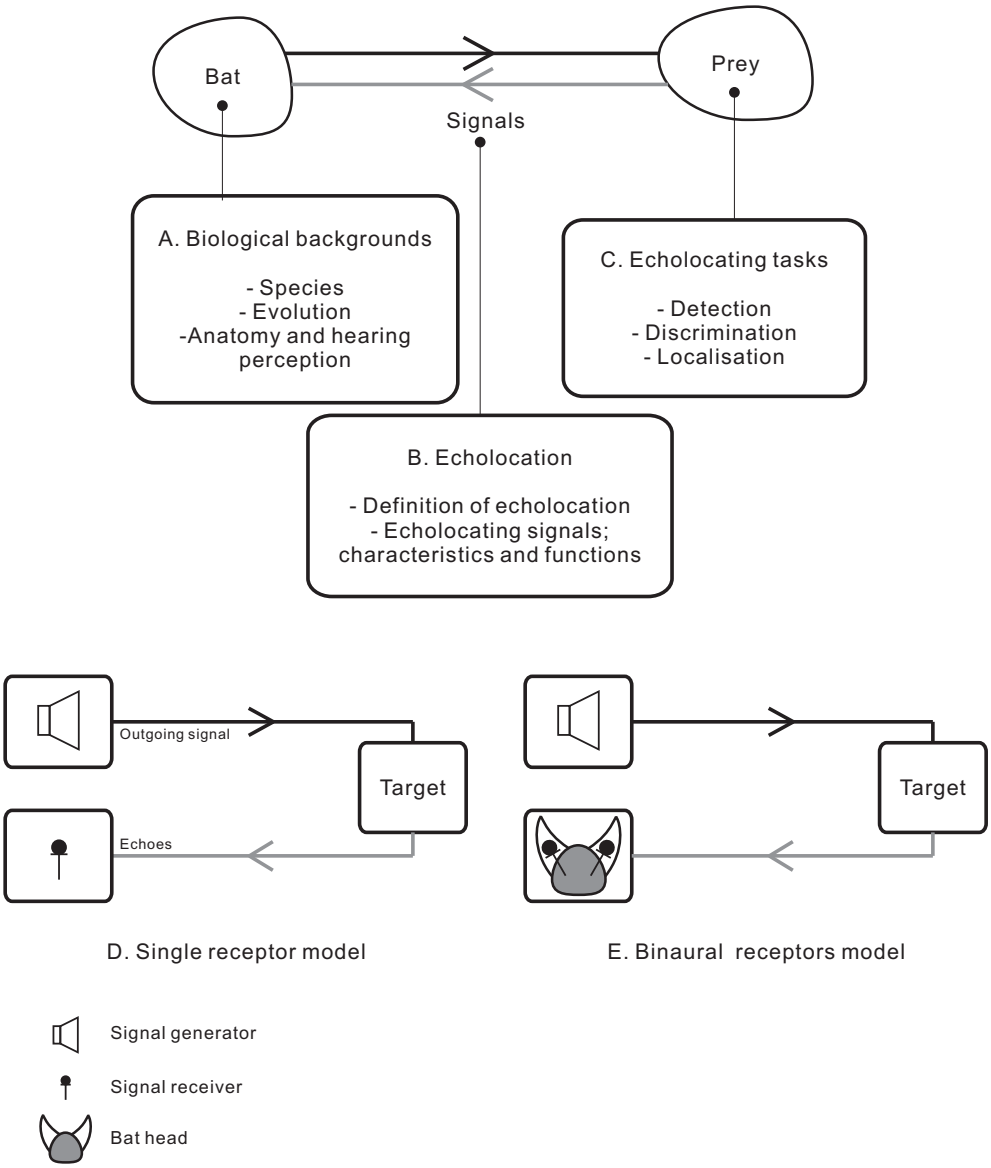


Figure 2.1: Main contents in the review structure

ears when the reflected signal is received. The single receptor model does not include these effects but the binaural receptor model considers them because two receivers at the bat's ears are modelled. The characteristics of each model will be described later in section 2.5. In section 2.6, a biol-inspired technology in relation to echolocation in bats is introduced. Finally a summary of the review and the aim of the current study will be described in section 2.7

2.2 Overview of bats

2.2.1 Species

Bats are mammals of the order Chiroptera. Their most distinguishing feature is that their forelimbs are developed as wings, making them the only mammal naturally capable of flight. There are estimated to be about 1,100 species of bats, accounting for about 20 % of all mammal species. About 70 % of bats are insectivorous. Of the remainder, most feed on fruits and their juices, three species sustain themselves with blood and some prey on vertebrates. All bats are included in the order Chiroptera (meaning hand-wing) which has two suborders, *Megachioptera* (Megabats) and *Microchiroptera* (Microbats). *Eptesiscus fuscus*, a species of Microbat, and *Rosettus aegyptiacus*, a species of Microbat, are shown in Figure 2.2.

(a) (b)

Figure 2.2: Bats (a) *Eptesiscus fuscus, Microchiroptera* (b) *Rousettus aegyptiacus, Megachioptera*

Microbats live on every continent except Antarctica, and all use echolocation for obstacle avoidance and foraging. They are generally smaller than Megabats, their ears are often large and complex and many species have nose leaves. These features are known to be associated with their echolocating ability. All species of the Microbats have advanced echolocation capabilities although they often have good vision. Their echolocation sounds are produced in the larynx (vocalisations). Most Microbats feed on insects, whereas most Megabats eat fruits or suck nectar from flowers. The Megabats inhabit tropical Africa and Asia and rely mainly on sophisticated vision and smell for flight and foraging. Except for one genus, *Rousettus aegyptiacus*, Megabats are not capable of echolocation. *Rousettus* lives in caves and uses a crude form of echolocation to avoid obstacles when in complete darkness. The echolocation calls that *Rousettus* uses are completely different from Microbats in terms of echolocating signal generation. They produce the sound by clicking the tongue against the side of the mouth.

2.2.2 Evolution

More than 150 bat species may coexist in ecological community, a number that far exceeds that of any other mammalian group. Despite their prominent position among mammals, the evolution of bats is not well known because of poor fossil records and incomplete evolutionary trees. Since the skeleton of a bat is small and delicate, it does not fossilize well.

Living bats are classified into 18 families on the basis of shared anatomical specialization and echolocation habits. To date, the study of echolocation is affected by the debate about the phylogeny of bats because most species of the Megabats do not echolocate. Generally, it has been assumed that echolocation evolved only once in bats because it is a complex system involving specialisation of the respiratory system, ear, and brain. This hypothesis has been supported by the idea that the origin of bats is basically split into two groups, one lineage is echolocating bats (Microbats), the other is non-echolocating bats (Megabats). However, a competing hypothesis is that echolocation evolved independently in both suborders and not in one lineage. Although the general relationships between the bat families appear to be clear, there is still no general agreement about whether the Megabats are the closest living relatives of the Microbats.

Recently, as shown in Figure 2.3, Teeling et al. [2005] provided a new evolutionary tree which helps to explain how, when and where bat species originated. The result from the study supports a hypothesis that megabats are nested among four major Microbat lineages. Teeling *et al.* speculate that bats diversified in the Early Eocene period (about 52 to 50 million years ago) in response to significant global rise in temperature, increase in plant diversity, abundance of prey in variety, and that varied echolocation and flight strategies which characterise the families may have evolved.

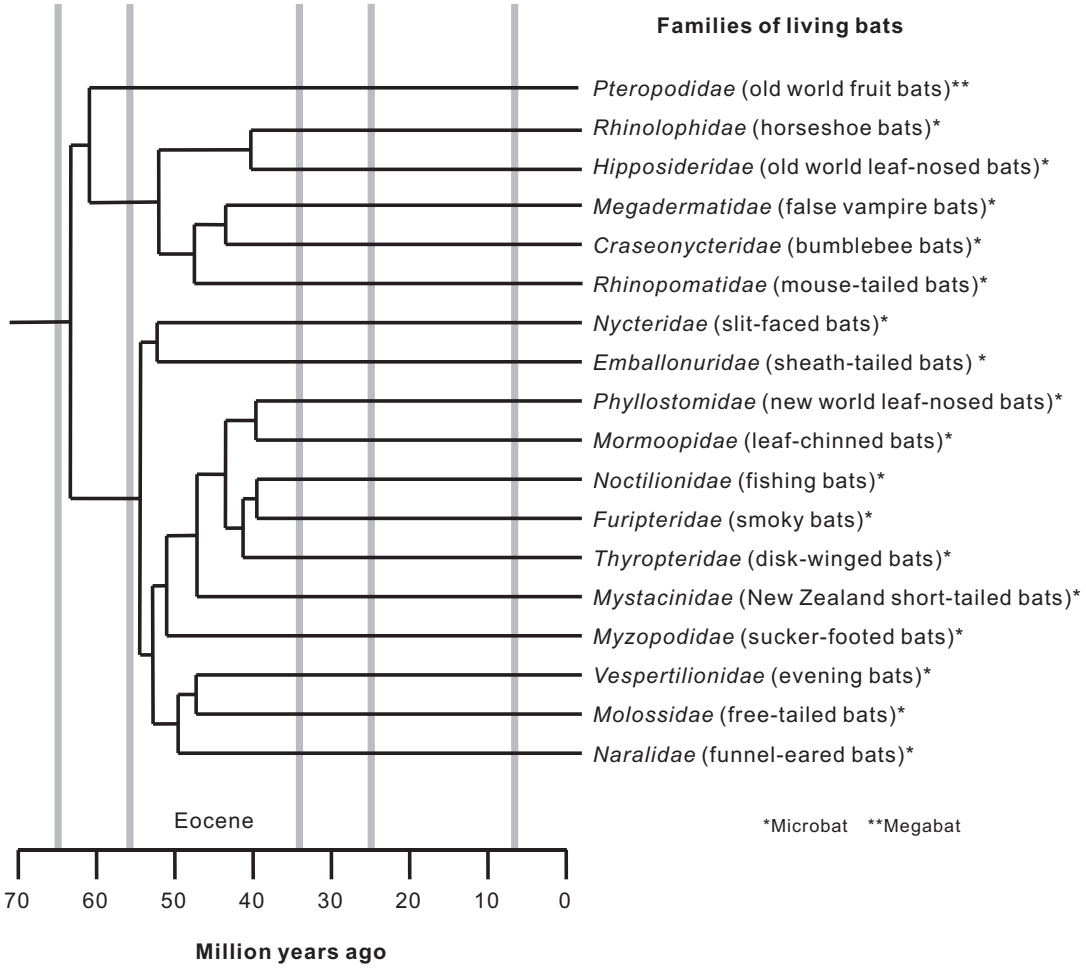


Figure 2.3: Temporal pattern and evolution of bats

2.2.3 Anatomy and hearing perception

An important reason for the evolution of the mammalian auditory system is to localise behaviourally relevant sounds. The shape of the outer ear varies widely across species but all mammals share similar anatomy and physiology in their auditory system. Without exception, the anatomies of bats basically resemble that of other mammals in form and function.

However, the outer ear of the bat is highly specialised. The pinna of bats are relatively large compared to their head size, or even many times bigger, and they have a flap-like shape. Various shapes of pinna are shown in Figure 2.4. The inside surface of the ear pinna frequently has several transverse ridges or a series of longitudinal ridges, which are presumed to provide structural support from the pinna. The tragus and anti-tragus are also part of the outer ear. The tragus is a small 'earlet' positioned in front of the ear canal and it has been known to have a considerable effect on vertical sound localisation as well as the pinna [Lawrence and Simmons, 1982a]. The anti-tragus is a broad flap that is continuous with the outer margin of the pinna [Hill and Smith, 1984].

The anatomy of the bat's ear is shown in Figure 2.5. In the middle ear of Microbats, the overall structure is similar to that of other mammals but the tympanic membrane, which is the thin partition transmitting the sound energy to the middle ear, is relatively thinner. Generally, it has been shown that bats which use high frequency ranging from 50 to 125 kHz have a smaller tympanic membrane than those which operate at lower frequencies. During echolocation the stapedius muscle of the middle ear plays an important role as a 'pulse and echo gain-controller' for the signal entering the cochlea. The contraction of the stapedius muscle strongly attenuates sensitivity during the echolocating pulse emission and weakly attenuates the response to echoes from the targets for the auditory system to maximally gain echoes [Grinnell, 1995]. This is important in bat's hearing because echo energy is very low compared with the emitted echolocating pulse.

In the inner ear, the cochlea of Microbats is specialised for the use of high frequencies. It has been noted that cochleae of Microbats are sensitive and sharply tuned at high frequencies up to 160 kHz. There is some physiological evidence in cochlea anatomy to support high frequency hearing in bats. The basilar membrane forms part of the Organ of corti and it houses the mechanism which plays the major role in transmitting sound waves and in frequency analysis. According to Békésy [1989], high frequency signals displace the basal cochlea region which has a high stiffness, and low frequency signals

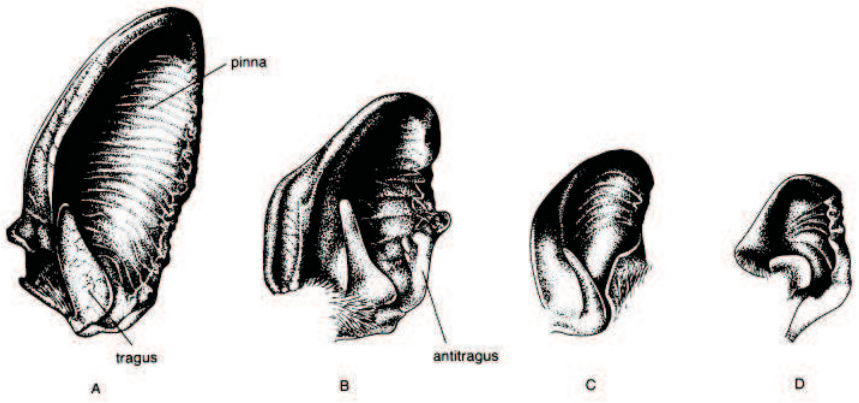


Figure 2.4: Pinnae of several Vespertilionidae (after Hill & Smith, 1984). (A) *Plecotus auritus*. (B) *Barbastella barbastrellus*. (C) *Myotis daubentonii*. (D) *Nyctalus noctula*.

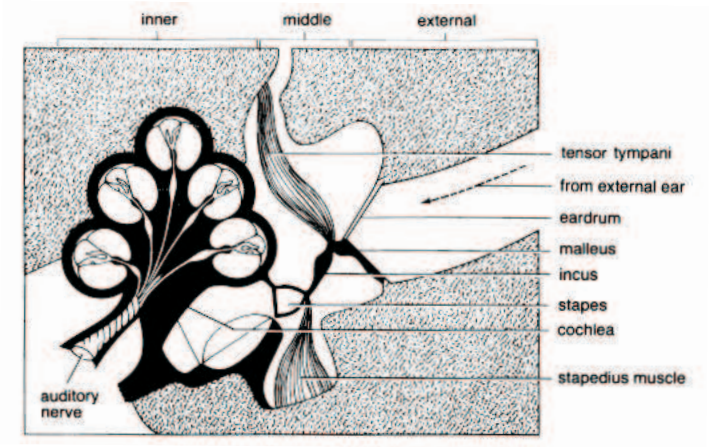


Figure 2.5: Anatomy of the bat's ear (after Hill & Smith, 1984)

displace the region which has a low stiffness. For Microbats, Kössl and Vater [1995] have shown that they have a narrower and thicker basilar membranes than those of non-echolocating mammals, which explains the adaptation to high frequency selectivity. Also, the hair cells, which play the role of signal receptor, are of shorter length in Microbats than those mostly presented in other mammals. This evidence corresponds to the extended high frequency hearing range of the bats as well.

An interesting and specialised feature of the cochlea is found in the CF/FM bats. Such bats are observed to have an *acoustic fovea* in the cochlea where the neurons are specifically tuned to a narrow frequency range of returning echoes and are seen in Doppler shift compensating bats such as *Rhinolophus ferrumequinum*, *Pteronotus parnellii* and *Asellia tridens* [Schnitzler and O. W. Henson, 1980]. The degree of sharpness in the tuning curve can be characterised by the Q_{10} dB values (centre frequency divided by

bandwidth 10 dB from the tip of the bandpass filter). The sharper the tuning, the higher the Q_{10} dB value which gives better frequency resolution and accuracy. For example, the mustached bat's cochlear is specialised in response to the first-harmonic of the CF component that it is able to measure the target velocity and detect the wing beats of insect prey [Neuweiler, 1990]. Its behaviour is supported by the evidence that the Q_{10} dB values in the tuning curve can reach 600 in that frequency region. The other example shows that the frequencies a few kHz above and below the carrier frequency can reach Q-value up to 400 [Kössl and Russell, 1995]. Another piece of biological evidence supporting the notion of the acoustic fovea is that the outer-hair cells (OHCs) of the fovea constitute 40 % of the total population of the cochlea and OHC responses to the carrier frequency signal dominate the response from the fovea [Russell et al., 2003]. Typically the response from the cochlea of the CF bat is known to continue to ring for several milliseconds immediately after the offset of tones at or close to the carrier frequency, but not at frequencies higher than 2 kHz above and below the resonance [Russell et al., 2003] .

The behavioural audiogram of several species of bats have been reported in previous studies as shown in Figure 2.6. An audiogram is a graph that shows the softest sounds one can hear at different frequencies. A frequency range where a bat responds to thresholds below 60 dB SPL is assumed to be within the bat's auditory range. From the literature, FM (frequency modulated) bats are not only able to hear higher frequencies, but also at extremely low levels as well. For example, the auditory threshold of *Megaderma lyra* remains close to 0 dB up to frequencies of about 100 kHz, and the threshold of hearing between frequencies of 16 and 45 kHz is below 0 dB. It is also interesting to see the notches and peaks in the audiograms. It can be explained by the observation of Doppler shift in echolocating signal as the bat moves relative to its prey. In the echolocation of CF/FM bats, the emitted echolocating pulse and returning echoes can overlap in time and, therefore, there is an obvious need to filter out the emitted signal. This is achieved by the Doppler-shift and sharp maxima and minima of the auditory threshold curve. The frequency of the emitted signal is actively tuned so that the returning echo of interest falls into the sharp threshold valley, while the frequency of the emitted pulse is in the insensitive region of threshold of hearing. These frequency characteristics are shown in most of the echolocating bats. It has to be also noted that the hypersensitive region of the frequency range is also used for passive listening and communication calls.

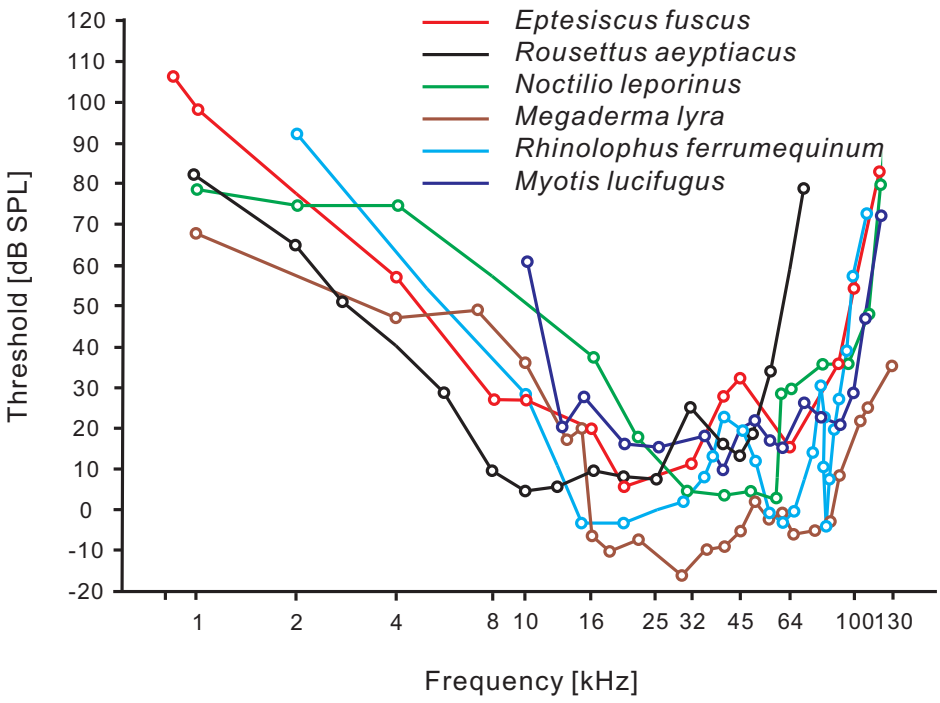


Figure 2.6: Behavioural audiograms in dBSPL for six species of bats: *Rousettus aeyptiacus* (Egyptian fruit bat; Koay *et al.*, 1998), *Eptesiscus fuscus* (big brown bat; Koay *et al.*, 1997), *Megaderma lyra* (Indian false vampire; Schmidt *et al.*, 1983), *Myotis lucifugus* (little brown bat; Dalland, 1965), *Noctilio leporinus* (fish-catching bat; Wenstrup, 1984), and *Rhinolophus ferrumequinum* (greater horseshoe bat; Long & Schnitzler, 1975) (after Koay *et al.*, 1998)

2.3 Echolocation

2.3.1 Definition

Echolocation usually refers to an acoustical and physiological process used by several mammals such as bats, dolphins and whales for locating and identifying invisible objects by emitting sound waves that are reflected back to the emitter from the objects. Echolocating animals emit their own pulses into the environment, and listen to the returning echoes from various objects in the environment. The term was coined by Griffin [1958], who was the first to conclusively demonstrate its existence in bats. It has evolved in many kinds of vertebrates in different forms. Simple forms of echolocation are used by cave-dwelling swiftlets and oilbirds and by small nocturnal mammals such as shrews, rats, and tenrecs. More sophisticated echolocation systems are found in bats and toothed whales. In particular, the latter groups use echolocating signals for navigation and for foraging in various environments. During echolocation, target ranging is done by measuring the time delay between the animal's own pulse emission and the echoes. The echoes returning to the two ears arrive at different times and different intensity levels, depending on the position of the target producing echoes. Spectral information in echoes reaching into two ears also varies depending on the target. Information in the returning echoes such as difference in time and intensity levels and the frequency spectrum is used in echolocation including target localisation and discrimination

Echolocating bats depend upon their biological pulse emission to detect, identify and locate their prey. The echolocation pulses they use consist of short bursts of sound at frequencies ranging from 10 kHz to about 200 kHz [Lawrence and Simmons, 1982b]. Signal information such as distance, range, angular position, shape and texture of an object are used in the process.

2.3.2 Echolocating signals

Microbats fall broadly into two categories depending on the types of echolocating signal they use; those that just use FM signals, and those that use a combination of FM and CF signals. The latter category are often referred to as CF bats. Each species of bat emits echolocating signals with different duration, frequency bandwidth and SPL (sound pressure level). The particular type of signal used in each species of bat enables effective echolocation for its own processing mechanism. Schnitzler and O. W. Henson [1980] demonstrated that narrowband, shallow FM components seem to be well suited

for the detection of echoes from small insects and target movements, whereas broadband, steep FM components seem to be well suited for accurate target localisation by delivering exact time markers. It is generally regarded that broadband echolocating signals carry better spectral information in the target reflection. On the other hand, CF components seem to be suitable for the Doppler shift compensation, in combination with CF bats' specialised hearing system and acoustic fovea. Therefore, the CF echolocating bats can classify fluttering insects by detecting temporal modulation patterns.

The observation that each different type of the signal has its own advantage and disadvantage corresponds the fact that the bats live in various conditions. Considering the different foraging environments, such as open or closed spaces, and the types of prey that bats echolocate, the evolution of diverse signal structures seems necessary. In other words, the characteristics of an echolocating call are adapted to the particular environment, hunting behaviour and the type of food source. Although there is difference between different types of signals, the general characteristics shown in the most of echolocating signals can be described as follows.

a. SPL The SPL of emitted signal can reach up 100 dB or more with air reference sound pressure 0.02 mPa. It has been recently reported that, at 10 cm distance from the bat, an intensity greater than 125 dB peak level has been shown in recording from insect eating bats and 133 dB has been shown in the recording from open space flying bats [Jones and Holderied, 2007]. It has been also reported that the echolocating calls from stationary bats were about 13 dB less intense than the calls during flight [Waters and Jones, 1995]. Most of the bats are also known to contract their middle-ear muscles in synchrony with emission of echolocating calls to prevent the damaging in the auditory system [Suga and Jen, 1975].

b. Duty cycle The duty cycle of a signal is defined as the ratio of the pulse duration (time) and the period between pulses, and is affected by the temporal pattern of signal delivery. In echolocation, the signal duty cycle can be maximised by increasing the signal duration, signal emission rate, or both. The duty cycle in echolocation is important in describing any overlapping pattern between the pulse and echo.

c. Frequency patterns In most of echolocating bats, the frequency modulated (FM) downward sweep and the constant frequency (CF) tones comprise the echolocating signal. The FM sweep is a broadband signal and the CF signal is a narrowband signal. Both of the signals can appear in one frequency range or in multiple harmonics.

d. Harmonics There exists a dominant harmonic component which has the highest intensity when an echolocating call is composed of multiple harmonics.

e. Duration A single continuous echolocating signal can last from 0.2 to 100 ms in duration. Typically these last several milliseconds and are emitted 2-10 times per a set of echolocating signal. After the bat detects a target of interest, it increases the repetition rate and decreases the duration of its signals to avoid overlapping between the emitted call and the returning echo.

f. Delay During the flight while the bat is emitting the signal and searching for the prey, a delay occurs between the call and the returning echoes. This is related to range detection and target separation.

As described above, each echolocating signal carries its own information. The target information from echolocation is encoded by returning echoes and their relationship with the emitted signal. The information obtained by the processing of the signals as follows.

a. Range Range is the distance from the target object and it can be calculated from a specific delay between the original call and the echo.

b. Size Information on size includes length, width, shape and thickness. It is represented in the waveform characteristics such as amplitude, peaks and notches in time domain.

c. Velocity Some echolocating bats, particularly CF-FM bats can calculate the relative velocity of a its target object. This is achieved by analysing the changes in echo frequency caused by Doppler shift and comparing them to the original echolocating call. The amount of Doppler shift is proportional to the relative velocity.

d. Fine structure This refers to the spectral characteristics of the returning echo from a specific target. For example, it might the edibility of various targets from the different echo structures.

e. 3-D location This is the localisation of the horizontal and vertical positions. The basic principle is same as human localisation [Blauert, 1997]. The horizontal position is determined by binaural hearing which compares the difference between the signals reaching the two ears. The vertical position is determined by using information varying

according to the pinna and tragus in bats. The external ears generate different spectral peaks and notches at different levels of elevation, giving vertical cues.

f. Flutter of target The insectivorous bats discriminate a fluttering insects with various wingbeat frequencies [Grossetete and Moss, 1998]. It has been shown that FM bats can discriminate the fluttering target at different frequencies by using the delays between echoes from the moving and stationary parts of the target [Vonderemde and Menne, 1989].

2.4 Echolocation task

The tasks that bats achieve for successful echolocation can be categorised as detection, discrimination and localisation. In this section, the behavioural evidence observed in the three different tasks are described.

2.4.1 Detection

Although most echolocating bats have a high sensitivity of hearing to their signals, they have to detect these signals against background noise such as wind, self-generated noise during flight, and communication signals produced by other animals. Echolocating sounds used by bats may exceed 100 dB SPL at a distance of 10 cm [Griffin, 1958]. Because of this intense sound level of vocalisation, the middle ear muscle of the bat contracts to protect the inner ear from damage. However, the attenuation effect due to absorption of high frequency sound propagating through the atmosphere has been shown to be from 0.7 dB/m at 30 kHz to 8 dB/m at 200 kHz [Lawrence and Simmons, 1982b]. Furthermore, Lawrence and Simmons reported that in a conditioned laboratory experiment using a 19 mm diameter spherical target at 3 m the attenuation and additional spherical spreading loss result in only around 10 dB SPL of returning echoes which is decreased by 100 dB compared to 110 dB at 10 cm from the sound emitting source. The schematic diagram of sound level attenuation in this echolocation example is shown in Figure 2.7. Therefore, the returning echolocating signals can be interpreted as relatively ‘very low level’ signals.

2.4.2 Discrimination

There have been a few studies about how bats discriminate objects. The discrimination strategy is known to be different depending on the type of echolocating signal

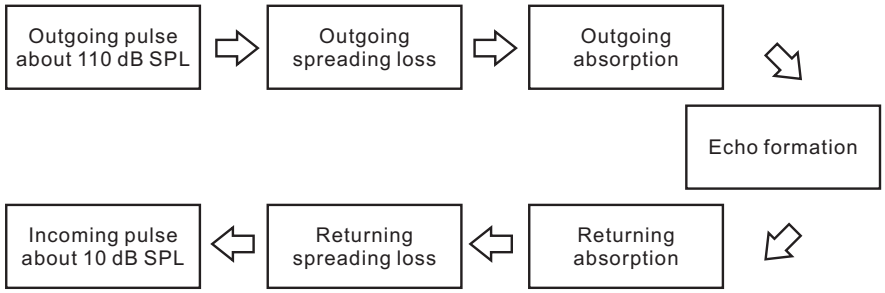


Figure 2.7: Schematic diagram of sound level attenuation in echolocation

used. Accordingly, the acoustic characteristics of targets also vary depending on the foraging type, habitats and wing morphology of prey. The bats that use a CF signal in echolocation utilise the velocity of the flying insect by detecting Doppler shifts. The echoes returning back from the prey when a CF signal is used, show the frequency rapidly modulated by the insect wing beats. Therefore those bats are able to track the movement of the insect. It has been suggested that FM bats also utilise the Doppler-induced changes in echoes to discriminate the fluttering target [Grossetete and Moss, 1998]. In addition, a previous study has suggested that FM bats are able to use an 'acoustical glint' of the target [Simmons et al., 1990]. The acoustic glint is a series of reflections from the complex target in the temporal domain. Simmons *et al.* examined the acoustic glint from mealworms, spheres and discs and the results were shown to provide the profile of them such as the shape information [Simmons et al., 1990]. The same study provided the behavioural evidence observed in FM echolocating bats that it showed they can discriminate edible mealworms from the inedible spheres of various sizes. The acoustical properties of those different targets showed the different range profile in terms of temporal spacing of the spectral peaks and notches [Simmons and Chen, 1989]. The acoustic glint might not be effectively generated when a CF signal is used because the CF signal is consists of a dominant frequency and its harmonics. Hence, it can be expected to carry less information in terms of spectral analysis than the broadband FM signal.

2.4.3 Localisation

It has been noted that good high frequency hearing is more common among mammals than good low frequency hearing [Heffner and Heffner, 1990]. This corresponds to the fact that the 'intensity (or level)-difference pathway' in the mammalian ascending auditory system is well established and very similar across almost all mammalian orders. In particular, echolocating bats can be said to have more specialised 'intensity-difference pathway' than 'time-difference pathway' because they use much higher frequencies than

most of other mammals. A primary goal of the bats' auditory system is to localise behaviourally relevant sounds in space. It has been generally shown that the interaural level differences (ILDs), caused by a small head (which,for example, is less than 1cm of diameter for the big brown bat), enable them to localise high frequency sounds. They can resolve the auditory image with high spatial precision because the head acts as an obstacle reflecting sound waves.

In terms of investigation of localisation ability in bats, there has been many studies on the big brown bat (*Eptesiscus fuscus*). Big brown bats are relatively large in size, from about 10 cm to 13 cm in length and 14 g to 18 g in weight. Although they are not small among the echolocating bats, their head size is mostly less than 1cm. Generally it is known that big brown bats can detect echoes from small objects at a distance of up to 5 m. This species uses broadband frequency modulated (FM) sweeps, ranging from 100 kHz down to 25 kHz with two or three harmonics. Those FM signals are well suited for spatial localisation. Therefore, the big brown bat is considered a good model for the study of echolocating bats. The localisation abilities in bats studied from behavioural experiments are summarized in Table 2.1. To measure the resolution of localisation in big brown bats, living bats must be trained and their behaviour evaluated in experimental conditions.

The azimuthal discrimination experiments by Simmons et al. [1983] and horizontal tracking experiments by Masters et al. [1995] report about 1.5° accuracy in the echolocating big brown bat. However, studying the same species of bat in a passive listening localisation task, Koay et al. [1998] reported that their threshold for discrimination in the horizontal plane is 14°. It is noted that there is a difference in the measurement methodology. Koay *et al.* used a conditioned avoidance procedure to measure left-right azimuthal sound source discrimination while Simmons *et al.* used the method of limits which measures the threshold when the subject cannot perceive the sound any more. Despite the influence of the differences in methodology, better acuity in echolocation than in passive location in the horizontal plane seems plausible. From those results, it can be argued that binaural hearing works more efficiently in echolocation rather than in passive localisation. Echolocation basically starts from the comparison between the emitting pulse and the echoes which have a certain arrival time on return. In a horizontal angle discrimination task experiment, Simmons *et al.* pointed out that the observed performance of the big brown bat at the task matches the performance to be expected if 0.5 μs acuity of perception of sonar echoes can be applied to the binaural echo arrival time perception. The big brown bat might use the interaural time difference and

reconstruct the acoustical image of the target by identifying phase differences from the specific target direction.

Table 2.1: List of studies of localisation ability in bats

Study	Type	Parameter	Species of Bat	Result
Simmons *et al.* (1983)	Echolocation	Horizontal resolution	*Eptesiscus fuscus*	1.5°
Matsters *et al.* (1995)	Echolocation	Horizontal resolution	*Eptesiscus fuscus*	1.5°
Lawrence and Simmons (1982)	Echolocation	Vertical resolution	*Eptesiscus fuscus*	3°
Wotton and Simmons (2000)	Echolocation	Vertical angle acuity	*Eptesiscus fuscus*	2.9°-4.1°
Koay *et al.* (1998)	Passive localisation	Horizontal threshold	*Eptesiscus fuscus*	14°
Fuzessery *et al.* (1993)	Passive localisation	Angular resolution	*Antrozous P.Pallidus*	1°

Even in a small bat with an inter-ear distance of less than 2 cm, the ILDs can reach 40 dB which covers a significant portion of the dynamic range of sound intensities that the mammalian auditory system encodes [Erulkar, 1972; Harnischfeger et al., 1985; Obrist et al., 1993]. Jen and Chen [1988] has also reported that the measured maximum ILD has been shown to be approximately 25 - 30 dB. Grothe [2000] has concluded that ILD coding does not challenge the general abilities of the mammalian auditory system and consequently are used for sound localisation whenever ILDs are available. On the other hand, at low frequencies, when the wavelength is equal to, or longer than, the inter-ear distance, the head does not reflect sound waves. Therefore, ILDs are not available any more to localise sound. For example, the free-tailed echolocating bat, *Tadarida brasiliensis* is reported to be not able to localise sound below around 7 - 10 kHz using ILDs [Grothe, 2000]. Moreover, the sensitive region in the audiogram of the bat hardly reaches such 'low' frequencies.

The sound source displacements from a frontal position produce a difference in the time of arrival of the sound between two ears because it takes longer for the sound to arrive at the ear that is further from the source. This delay is referred to as interaural time difference (ITD). It has been generally assumed that bats do not use interaural time differences (ITDs) as a sound localisation cue. However, a few previous studies

have measured the ITDs and suggest the possibility of utilisation of such cues. Koay et al. [1998] estimated the maximum ITD experienced by the big brown bat during the measurement of passive sound localisation ability which appeared to be approximately 55 μs. Jen and Chen [1988] also estimated that the maximum ITD usually occurs at 80 °-90 ° in azimuth.

It has been reported that the spatial information is not uniformly distributed over all possible positions. Noticeably, the ILDs of several bat species are quite linear with azimuth although these characteristics are only shown over a limited range of about ±20 °.

2.5 Modelling

As reviewed in previous sections, the echolocation phenomenon in bats can be regarded as a mixture of complex problems which is associated with the signal processing of biosonar acoustics. More recently, the echolocation in bats has been inspiring engineers and more importance has been given to learning and applying their strategies to sonar systems. To reduce the complexity and investigate a single problem one is interested in, a model is useful as it can provide new parameters or limit a given parameter which controls the part of the model. Papadopoulos and Allen [2007] have recently proposed an acoustical model of the echolocation process in bats. The concept of the model is that each stage of echolocation, the signal emission, reflection and reception, is modelled as an individual section and the echolocating process is recreated by combining those different parts in series which is mathematically called 'convolution'. Convolution is defined as an operation on two functions, producing a third modified version of original function. The convolution can be used to calculate the response of a system to arbitrary inputs by using their impulse responses. Convolving two signals in the time domain equally means that their spectra are multiplied in the frequency domain. The convolution technique is widely used in different areas of signal processing. A similar application of the convolution technique is shown in digital audio processing which recreates virtual sound by convolving the anechoic source with the impulse response of a given room. The advantage of this technique is that it is not necessary to have the source in the room but only the impulse response of the room is required when the source is provided. When it comes to echolocation in bats, it is difficult to measure the echoes from the target when a real bat generates sound during its flight. But using the convolution technique, the same echoes can be recreated theoretically by acquiring the impulse response of the target and the bat's echolocating signal separately.

Figure 2.1, D and E shows the simplified concept of the model. There are two types of models. One is single receptor and the other is binaural receptors.

2.5.1 Single receptor model

Figure 2.1, D show the single receptor model. This model consists of the echolocating sound source, target reflection and single receptor. The model would reproduce and predict the signal received in front of the middle of the bat's head in echolocation for a specific target. As the reproduced signal contains the information such as amplitude, temporal characteristics and spectral characteristics, the target detection and discrim-

ination are considered to be available in this modelling.

2.5.2 Binaural receptors model

However, as reviewed earlier, the localisation task which the bat performs is only available when the binaural hearing is enabled. The function of the external ear and the head is essential to model the localisation of the target. Figure 2.1, E shows the binaural receptors model. It basically extends from the single receptor model and it performs binaural processing for the signal received in the single receptor. The information for binaural processing in different location of the sound source is conventionally known to be completely encoded in the transfer function of the head and external ears. It is commonly referred to as the head-related transfer function (HRTF, H_L-left ear, H_R-right ear), and its time (t)-domain version is the head-related impulse response (HRIR, h_L-left ear, h_R-right ear). Practically, the HRIR is measured at both ears in different azimuth (θ) and elevation (φ). It is then transformed into the frequency (f)-domain by the Fourier transform as

$$H_L(f,\theta,\varphi) \longleftrightarrow h_L(t,\theta,\varphi) \tag{2.1}$$

$$H_R(f,\theta,\varphi) \longleftrightarrow h_R(t,\theta,\varphi) \tag{2.2}$$

The amplitude and the phase is described at each ear to provide level and timing information as

$$A_L(f) = 20log_{10}|H_L(f)| \qquad \Phi_L = \angle H_L(f) \tag{2.3}$$

$$A_R(f) = 20log_{10}|H_R(f)| \qquad \Phi_R = \angle H_R(f) \tag{2.4}$$

The interaural level difference and interaural time difference can be obtained from the interaural transfer function which is represented from the left HRTF and right HRTF as

$$H(f) = \frac{H_R(f)}{H_L(f)} \tag{2.5}$$

Therefore, the ILD is obtained from

$$A(f) = 20log_{10}|H(f)| = 20log_{10}|\frac{H_R(f)}{H_L(f)}| = A_R(f) - A_L(f) \tag{2.6}$$

And the ITD is obtained from the group delay[1] (τ_G) information in the interaural

[1]The rate of change of the total phase shift with respect to angular frequency

transfer function.

$$\tau_G = \frac{1}{2\pi}\frac{\partial\Phi(f)}{\partial f} = \frac{1}{2\pi}\frac{\partial(\Phi_R(f) - \Phi_L(f))}{\partial f} \tag{2.7}$$

However, it is difficult to relate the group delays at high frequencies to the physiological functions as there is no evidence that the auditory system can resolve the high-frequency phase spectrum. For humans, since the maximum possible ITD is about 660 μs for a typical radius of human head (approximately 87.5 mm), the ITDs are generally useful only for the frequencies below 1.5 kHz. If the size of the bat is assumed to be ten times smaller than that of human, the ITD is expected not to be useful below 15 kHz, based on the scaling law. Therefore one typical method of modelling the auditory system is as a series of bandpass filters with a range of centre frequencies. For example, octave bandpass filters, which vary the bandwidth depending on the performance of the auditory processing, can be applied and they employ wider bandwidths for higher frequencies. Another alternative method is to compute the envelopes of the signals measured at the two ears, then cross-correlate the envelopes to measure the interaural delay. The envelope of a signal is the outline of the signal and it connects all of the peaks in the signal. Using the envelope delays, one can obtain the interaural delays which increase monotonically with the angle of incidence of the sound relative to the median plane. Extracting an envelope is useful for the analysis of interaural delay in bats as the ITDs are not available for most of the high frequencies in the echolocating signal.

For a symmetrical head, both ILD and ITD are zero in the median plane. Generally, it is accepted that the measured HRTF is a rather complicated joint function of azimuth and elevation. There are also substantial inter-subject differences in frequency response. Although the HRTF shows limited systematic variation of frequency response as a function of azimuth and elevation, HRIR measurement and HRTF analysis is required to reproduce and characterise the binaural signals which reach the eardrum in each ear.

In conclusion, the HRIRs can be convolved with the signal which is received at the single receptor. This enables simulation of the binaural echolocating signals used in localisation.

2.6 Binaural processing and bio-inspired technology

To develop the binaural receptors in the echolocation model, it is necessary to understand the characteristics observed in the HRTF of the bat-head cast. The bat-head cast is going to be used as a proto-type as this is regarded to be appropriate to simulate the presence of head and pinna. Using this proto-type, the current study aims to contribute to the knowledge supporting the bio-inspired engineering applications by providing insights into the sonar system using binaural processing. The binaural receptors have various potential future applications as they enable us to simulate not only object discrimination but also the localisation task in both horizontal and vertical axes. Note that the term "binaural receptors" implies two receivers (one at each ear) and it does not simply mean using double receivers specifically. It has been often misunderstood that binaural processing refers to using two receivers. Here, the terminology of binaural receptors will include both two receivers in conjunction with the HRTF and also double receivers for convenience.

It is also required to understand what has been found in terms of localisation in bats and what has been achieved so far in the bio-inspired application of binaural receptors. The former is mostly investigated in biology and the latter is investigated in engineering such as robotics, electrical engineering, computer science and acoustics. In biology, a few studies have measured the HRTF in bats and moreover, have shown that binaural cues vary systematically in both azimuth and elevation [Fuzessery, 1996]. The pinna of the bat create systematic changes in the spectral features of transfer functions and they perform an important role in localisation of targets in different elevations [Jen and Chen, 1988]. It is also known that the movements of pinna, which rotate in anti-phase are important for vertical localisation [Mogdans et al., 1988]. In the engineering field, the application which tracks a fixed or moving object has been developed by using the echo arrival times detected at two receivers [Barshan and Kuc, 1992]. Also, two moving receivers with a built-in robot sensor were inspired by pinna movements in the bat and the receivers were able to improve the localisation ability using a tone signal [Kuc, 1996]. These previous studies have employed narrowband signals and utilised the binaural cues such as ITDs and ILDs. However, there have been no studies which have adopted the artificial head and investigated the HRTF using broadband signals in the sonar system. On the other hand, a human-inspired binaural mobile robot which localise the sound has been designed using a neurological auditory processing which adopts general mammalian auditory model [Liu et al., 2008]. Hence, a biologically-inspired system is not only limited to a specific animal and there exists substantially similar mechanism across the mammalian binaural hearing which motivates the devel-

opment of biological system. The bat-inspired binaural system can be described as a development of echo-specialised processing based on the general mammalian binaural hearing. In fact, the bat-specialised features such as types of pinna has been highly motivating the related research. However, according to the reviewing article about biomimetic sonar system, the effect of binaural hearing on the echolocation which includes the presence of the head has rarely been investigated in the previous studies although the two receivers were employed in the robotic system [Müller and Kuc, 2007].

In addition, in previous studies of biology in bats, the HRTF of real bats have been measured at a fixed distance which varied from 15 cm to 1 m in different studies. Therefore, there seems to be no standardised distance for the measurement of HRTF in bats. Arguably, the movement along the distance is another parameter which may contribute to the echolocation. The bats fly and move towards the target when capturing their prey. The key question is which information they obtain as they approach the target, how the information changes, and whether the change in the information contributes to capturing prey. Therefore, the HRTF of the bat at various distances motivate current study.

2.7 Summary and aims

It has been shown that there are different ways of examining the echolocating abilities of bats. The achievement of echolocating tasks seems to be more complicated because it is associated with the acoustical transmission of the echoes from a specific target. It is suggested that each stage of the echolocation processing can be modelled by a mathematical function of convolution and that these can be combined together to simulate the given echolocating task. In the first place, one of the aims of this study is the testing of the single receptor echolocating model. Chapter 3 includes the following.

- Application of the model in object discrimination and investigation of the performance of various types of CF/FM signals.
- Extension of the object discrimination model to the auditory processing.

In this study, the experimental data from target discs is implemented with the model and the echolocating signals are artificially generated, therefore, neither living bats nor behavioural testing are involved.

Secondly, the current study aims to eventually investigate the binaural receptors model. To do so, data for a bat HRTF must be obtained. As there is no standardisation of HRTF measurement in small-headed animals, this study reviews the previous HRTF measurement studies in small-head animals, mostly bats. Then, in Chapter 4, the HRTF is measured using a bat-head cast. A soft bat-head cast, which was taken from a real bat-head, is used in this study. The analysis of the measured HRTFs is intended to generalise the HRTF characteristics of the bat-head cast rather than to investigate the species-dependent features of a particular bat. Chapter 4 includes the following.

- Examination of the repeatability in the measurement of HRTFs of the soft bat-head cast.
- Analysis of the species-independent characteristics of HRTFs of the soft bat-head cast.

In Chapter 5, a hard bat-head cast was made using the soft cast to prevent shape deformation. Then, the HRTF of the hard bat-head cast are measured at various distances. The measurement are conducted with the same set-up which has been built in Chapter 4. The distances are varied from distant (over 0.5 m) to nearby (less than 0.2 m) sources and the positions of the sound sources are varied in all azimuths with5 ° of increments at 0 ° of elevation. The measurement in different elevation are not carried out in this study. The analysis in the time and frequency domains are carried out to extract information of the measured HRTFs at different azimuths and the effect of the distance are demonstrated. The results finally provide the profile of the binaural receptors which is experimentally obtained.

In addition, a theoretical investigation is conducted by using a computational model which predicts the HRTF of a rigid sphere. Rabinowitz et al. [1993] have suggested a mathematical solution which predicts the sound pressure levels at the surface of the sphere. Then Duda and Martens [1998] have developed the computational model to show the human HRTF when sound sources are located at different distances. As the principle of the HRTF is similar between bats and humans in terms of the function of head and two ears, the previous studies of humans provide a good resource to investigate the HRTF of the bat head. Using the algorithm suggested and provided by the previous studies, this study predicts the HRTF of the sphere which has similar size to the bat-head cast. The results are expected to provide insights into the effect of the presence of non-sphere shaped head and pinna in the sonar system. Also the comparison between the theoretical results and the measured data are considered to be available. Therefore, Chapter 5 includes the following.

- Validation of the measurement system.
- Investigation of the effect of varying distances on the HRTFs of the hard bat-head cast.
- Simulation of the sphere model and comparison between the measured and the simulated data.

In Chapter 6, the general discussion is made to demonstrate the implication of this research and potential issues in designing binaural receivers in the sonar application based on the findings. In this thesis, the object discrimination tasks have been successfully applied to the single receptor model and it has also shown the effect of biological auditory system. Using the binaural receptors model, the implications for the object localisation task at different distances have been discussed. The results are expected to show how the localisation performance can be affected by the bat head when it approaches the sound source.

Chapter 3

Monaural Modelling on Object Discrimination

This chapter focuses on the single receptor model. First, the air transmission model simulates the process of the emission, reflection and reception at a single receptor. Then, a hearing model is designed to perform the auditory processing of the received signal from the air transmission model. Section 3.1 describes the air transmission model and its applications. Section 3.2 describes how the auditory model of the bat was designed and presents results from object discrimination studies.

3.1 Air transmission model

3.1.1 Introduction

The air transmission model describes the process of echolocation from the emission of the bat's signal to the reception of the returning echo signal. The physical quantities involved in the problem at hand are described in Figure 3.1 The bat emits an echolocation signal for which an acoustic pressure disturbance p_{out} is measured at point $\mathbf{r}$ at a distance r and azimuth (θ) and elevation (ϕ) relative to the centre of the interaural axis in the bat's head. The emitted signal propagates through the air until it is backscattered by the target creating a pressure disturbance p_{echo} back at point r.

The model described above enables us to separate the echolocating sequences involved and simulate each stage assuming a linear processing. Therefore one can predict the echolocating signal (p_{echo}) from a particular target that reaches the point before diffraction by the head based on the model. It is achieved by convolution of the emitted signal (p_{out}) with the backscattered response of the object ($h_{backscatter}$) at the orientation and distance shown in Eq 3.1

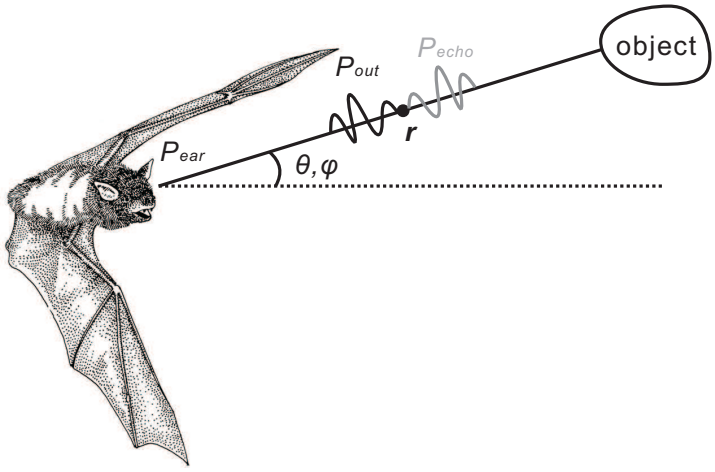


Figure 3.1: Diagram of echolocation system

$$p_{echo}(n) = h_{backscatter}(orientation, distance, n) * p_{out}(n) \tag{3.1}$$

where n denotes the sample number in discrete-time notation. In this study, both constant frequency (CF) and frequency modulated (FM) signals which represent the bat's echolocation signal are used to model the outgoing signal. The echo signal is simulated by convolving the outgoing signal with the experimentally measured backscattering impulse response from different types of target objects of various dimensions and materials.

3.1.2 Measurement

The backscattering impulse responses of three different sizes of plastic discs and two different thicknesses of wooden discs were measured in the anechoic chamber. Table 3.1 describes the physical characteristics of the discs being used and the abbreviations used for each disc. The experimental set-up has been described in a previous study by Papadopoulos and Allen [2007]. The backscattering impulse response data was provided by a colleague and the author is grateful for the input of Dr T Papadopoulos in this. A brief description of the experimental set-up is as follows. The sampling rate was set to 160 kHz with the anti-aliasing filter set to cutoff at 40 kHz, hence the system was frequency-limited at this range. The distance between the centre of each disc and the microphone was set to 15 cm. The flat surface of each disc was positioned to face toward the speaker. Therefore, the azimuth (θ) and elevation (ϕ) were both set to 0 in the model. As each recordings of raw measured data contains the influence from the microphone and speaker system, this effect was cancelled out using the response measured at the same position without the disc (which is described as the free field

response). The equalised impulse response is generated by deconvolving the free field response from the raw data. This was achieved in the frequency domain by dividing the raw frequency spectrum by the free field frequency spectrum. The raw and equalised impulse responses of the LP disc is shown in Figure 3.2. The pulses corresponding to the emitted and reflected signals appear to be within a time separation of less than 1.5 ms which corresponds to 240 samples. Only the reflected portion was extracted from the pulse for the simulation. Thus 200 samples around the peak of each reflected signal were processed. The maximum peaks of the reflected signals were detected and 200 samples from each measured impulse response consist of 50 samples before the maximum peak and 150 samples after the peak. This synchronises the response of each simulation. Finally, each response was zeroed for 800 samples which corresponds to 5 ms, before and after the response. The zeroing process suppresses the unwanted noise and reflections appeared in the backscattering impulse response in time-domain.

Table 3.1: Disc specification

Material	Category	Diameter [mm]	Thickness [mm]
Plastic	Large (LP)	37	20
	Medium (MP)	30	7
	Small (SP)	15	2
Wood	Thick (TW)	64	20
	Thin(tW)	64	10

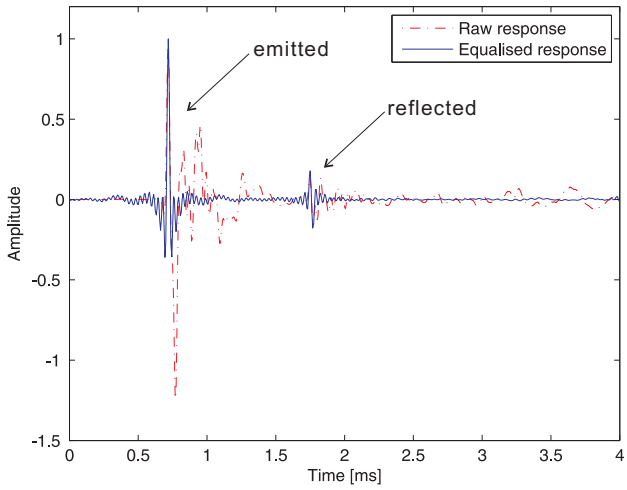
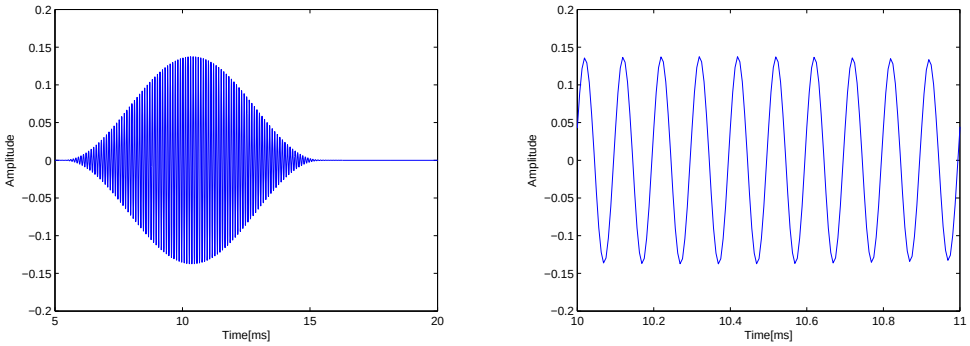


Figure 3.2: Impulse response of large plastic (LP) disc: raw response and equalised response.

3.1.3 Reconstruction and analysis of the signals

The ability of the echolocating signal to differentiate between experimental target objects was investigated by reconstructing the echoes when both CF or FM signals were used. The reconstruction and signal analysis procedure consists of two parts. One is the simulation of the echolocating signal and the other is spectrogram analysis. The simulation generates the echoes for a specific target and the spectrogram provides the frequency information along with echo generation time. A 10 kHz pure sinusoidal wave of 10 ms duration was chosen to simulate the CF signal within the frequency range available in the current system. There are many varieties of CF signals which are observed in different species of bats and these signals range from about 10 kHz to over 200 kHz in frequency and 0.5 ms to 50 ms in duration [Obrist et al., 1993]. In this study, a simple design of CF signal was selected for the simulation by the author. On the other hand, the FM signals are observed to have a diverse frequency bandwidth. The low frequencies range from about 10 kHz to over 180 kHz and the high frequencies from about 20 kHz to over 200 kHz, with various durations like the CF signals. Therefore, parameters for the FM signals of 10 ms duration were also selected by the author, with variations in sweep ratio (defined by the ratio between bandwidth and pulse duration). The 'chirp' function in Matlab was used to generate a linear swept-frequency cosine signal representing the FM signal. By changing the sweeping ratio, the bandwidth of each FM signal was set to 30 kHz-10 kHz, 25 kHz-10 kHz and 20 kHz-10 kHz each with the same 10 ms duration. Each artificially generated signal was processed by a 200-point Hanning window to avoid unnecessary ringing at the start and end edges of the frequency spectrum of the signal. Such ringing can occur due to the truncation in the temporal domain. The signals were then convolved with each backscattering impulse responses from the discs and each spectrogram was computed to produce the magnitude response in the frequency domain along the time axis. For CF signals, as described in the 'spectrogram' function in Matlab, the 256-point short-time Fourier transform and the same size of windowing in the time domain were applied with 128-samples overlapping. The phase response is not considered in this study as the spectral characteristics rather than the temporal characteristics are of interest. The convolved signal which reconstructs the echo from a CF signal on the LP disc is shown in Figure 3.3(a), and Figure 3.3(b) while Figure 3.4 shows the spectrogram of the reconstructed echo for the same disc. The overall amplitude pattern of the convolved signal in the time domain appears to have the shape of the Hanning window which was applied earlier but the signal looks like single tone sine wave when enlarged. The spectrogram shows a single smooth peak amplitude where the centre of the Hanning window was placed and does not appear to have multiple peaks and notches.



(a) Convolved LP signal

(b) Convolved LP signal, enlarged

Figure 3.3: Reconstructed CF signal (10 ms) from the LP disc (a) convolved LP signal. The envelope appears to have the shape of the Hanning window (b) the enlarged signal looks like single tone sine wave.

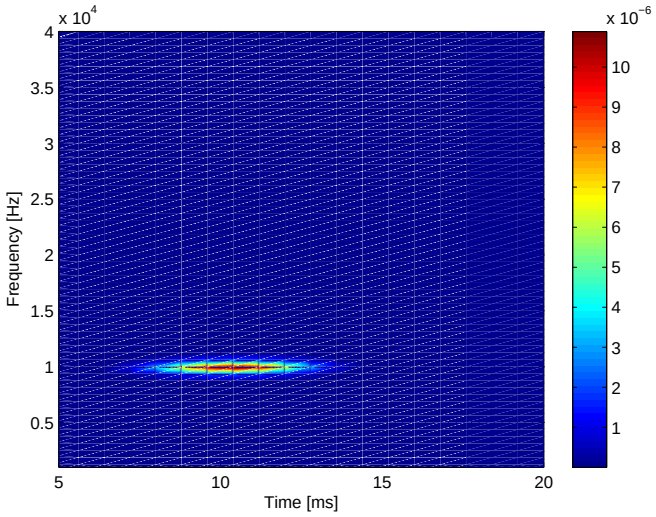


Figure 3.4: LP spectrogram. The power spectral density (PSD) at each frequency is plotted over the time sequence.

On the other hand, for FM signals, the size of the FFT and the degree of overlapping affects the representation of the spectrogram significantly due to the resolution in the frequency and time domains. Here, a 512-point short time Fourier transform with a Hamming window was applied to FM signals to increase the resolution in frequency and various degrees of overlapping were tested to find an optimum value for the processing. Figure 3.5 shows the time-domain signal of a FM chirp with a bandwidth 30 kHz - 10 kHz, and its spectrograms when the size of overlapping was set to 50 %, 75 % and 95 %, which correspond to 256, 384 and 486 samples of data. The result demonstrates that small amounts of overlapping generate unwanted artefacts in the spectrogram which are represented as several peaks in the PSD values for a linear chirp. This is considered as contouring error due to the mismatch between time and frequency resolution. As the extent of overlapping increases, the time resolution also increases and a single peak appears. Therefore, a 95 % overlap was used for the analysis of FM signals.

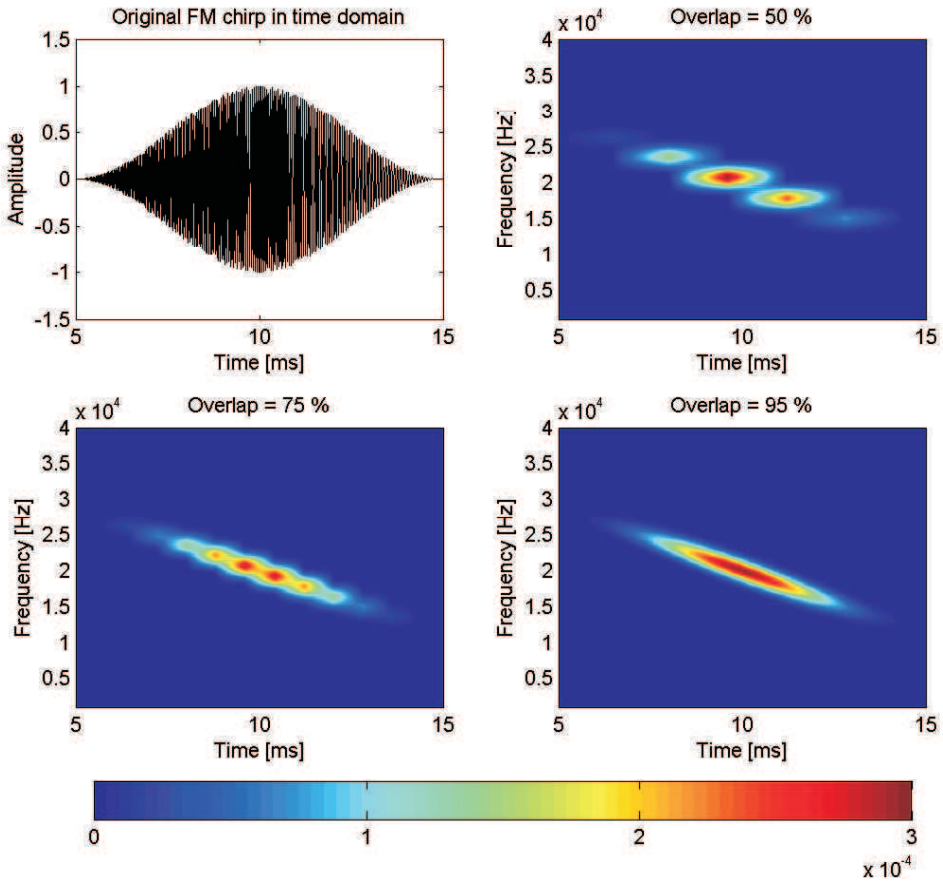


Figure 3.5: PSD spectrograms of original FM chirp (30 kHz - 10 kHz) and the effect of signal processing with various amounts of overlapping: 50 %, 75 % and 95 %.

3.1.4 Results

(1) CF signals

For CF signals, the magnitude of the power spectral density (PSD) at 10 kHz frequency bin was extracted from each disc's spectrogram. This was plotted along the time axis in dB as shown in Figure 3.6. The results show that both the dimension and thickness of each disc affect the magnitude of the echoes in CF signals. The TW disc, which is the largest and thickest of the experimental targets, showed the highest amplitude. This was followed by the tW disc which was approximately 13 dB lower in peak amplitude. The TW disc is thicker than tW by 10 mm though they have same diameter. On the other hand, the LP disc which has approximately half the diameter of the TW disc but which has the same thickness, showed about 28 dB of reduction in magnitude. The MP disc showed a 10 dB reduction compared to the LP disc which has a similar diameter but almost one-third of thickness of the LP disc. The SP disc which has half the diameter and less than one-third of the thickness of the MP disc showed about a 22 dB decrease in magnitude. These results suggest that the influence of the difference in diameter is larger than that of thickness. Additionally, because the thinner disc is slightly further away during the measurement, the amplitude of the echo is smaller. It is difficult to describe the material differences from the sample targets we have examined. It may be due to the lack of sample varieties to extract the parameter for the material. On the other hand, it may also due to the limitation of the CF signal to represent the object characteristics in the frequency domain.

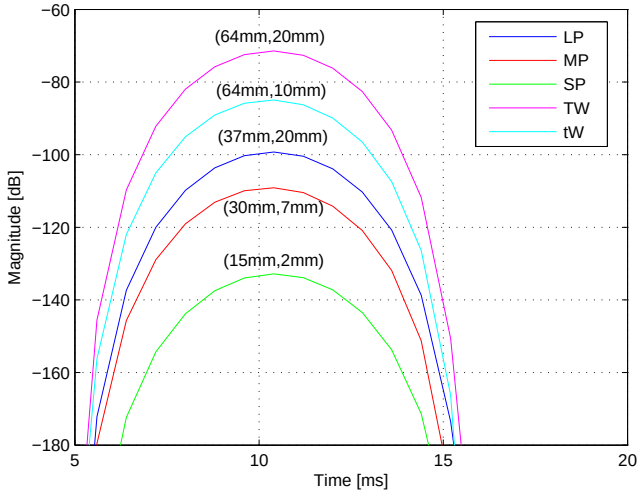


Figure 3.6: Analysis of CF echo signals, indicated as (diameter,thickness). The magnitude of PSD at 10 kHz frequency bin was extracted from each disc's spectrogram.

(2) FM signals

The effects of using FM signals of various bandwidths and sweep ratios were investigated. For the LP and TW discs, the effect of different sweep ratios was investigated. The analysis is conducted in terms of spectrogram's power spectral density (PSD). Note that the spectrograms of MP, SP and tW discs are also analysed in this study but they are not shown in the figure. The effect of various sweep ratio were appeared to be same between the discs, therefore, LP and TW are represented as examples. The spectrograms for the LP and TW discs for different sweep ratios are shown in Figure 3.7 and 3.8. Results from both discs appear to show more prominent peak and notch characteristics as the higher sweeping ratios (the larger bandwidth) are applied. In order to analyse the spectrogram information, maximum value of the PSD in each temporal occurrence was extracted and plotted. The schematic description of 'maximum PSD in temporal occurrence' is shown in Figure 3.9. Figure 3.10 shows the result of 'maximum PSD' for the original FM signal which the bandwidth is set to 30 kHz - 10 kHz. This shows the distribution of energy when the FM signal sweeps over the given range. The results of the simulated signals from the LP and MP discs are shown in Figure 3.11(a) and Figure 3.11(b). The analysis from the TW and tW discs are plotted in Figure 3.11(c) and Figure 3.11(d). From the graphs, two major properties can be identified. More prominent peaks appear when the 25 kHz- 10 kHz bandwidth signal is used than for the 20 kHz-10 kHz signal, but not necessarily with the 30 kHz- 10 kHz signal is used. The centroid of the PSD energy tends to move towards the earlier temporal occurrence when the 25 kHz-10 kHz and 30 kHz - 10 kHz signals are used more than when the 20 kHz - 10 kHz signal is used.

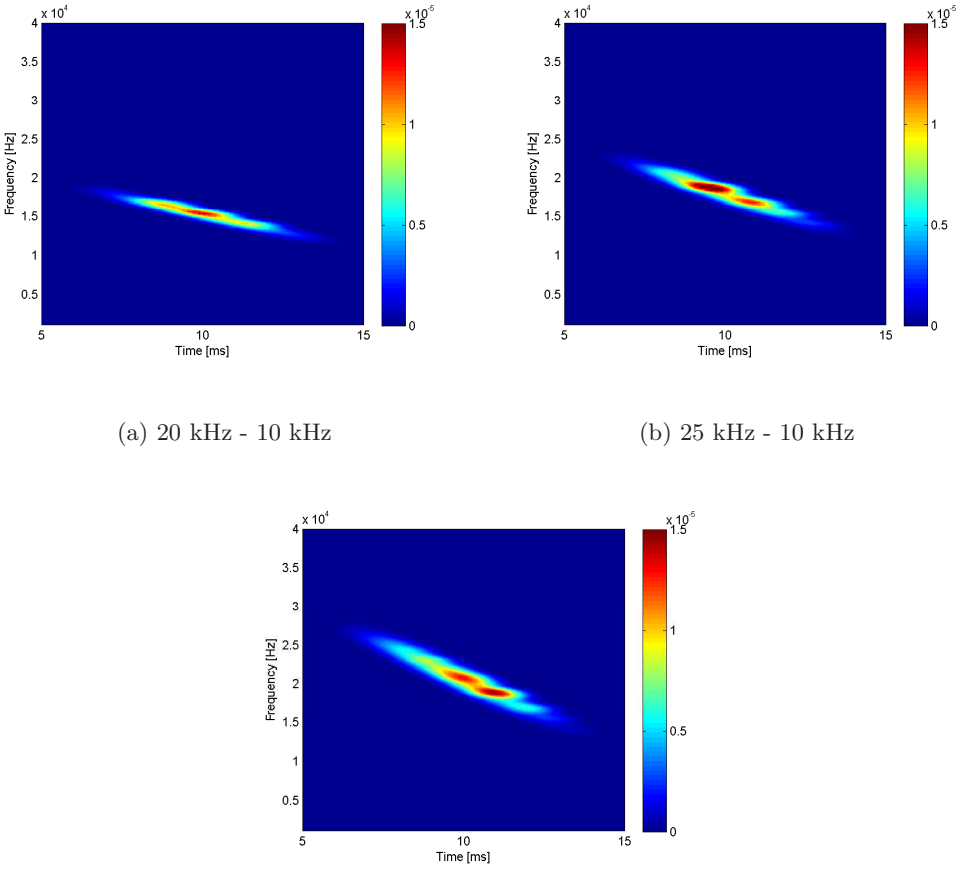
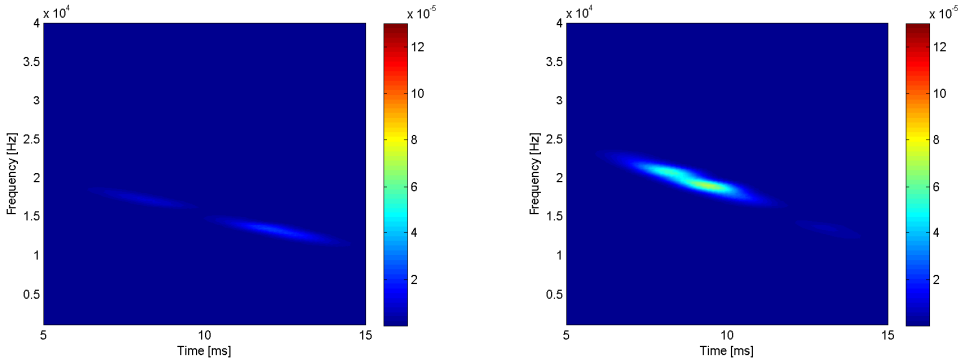
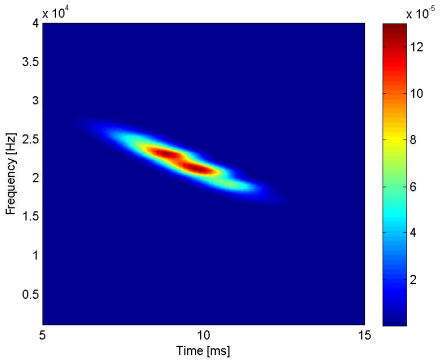


Figure 3.7: Power spectral density (PSD) from LP disc. The results are shown for the simulated echoes using 3 different types of FM swept signals. (a) 20 kHz - 10 kHz (b) 25 kHz - 10 kHz (c) 30 kHz - 10 kHz



(a) 20 kHz - 10 kHz

(b) 25 kHz - 10 kHz



(c) 30 kHz - 10 kHz

Figure 3.8: Power spectral density (PSD) from TW disc. The results are shown for the simulated echoes using 3 different types of FM swept signals. (a) 20 kHz - 10 kHz (b) 25 kHz - 10 kHz (c) 30 kHz - 10 kHz

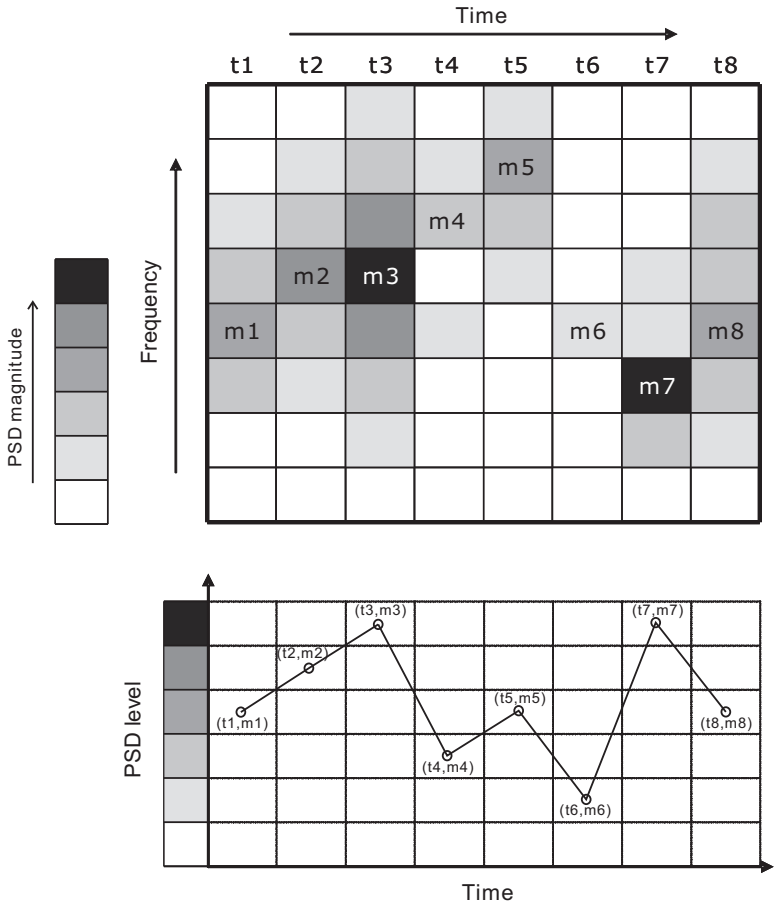


Figure 3.9: Schematic description of 'maximum PSD in temporal occurrence'. This shows an example of spectrogram (top) and its maximum PSD plot against time (bottom). t1-t8 denote time bins in analysis and m1-m8 denote the maximum PSD value in each time bin.

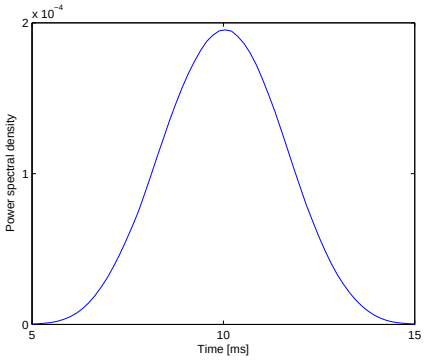
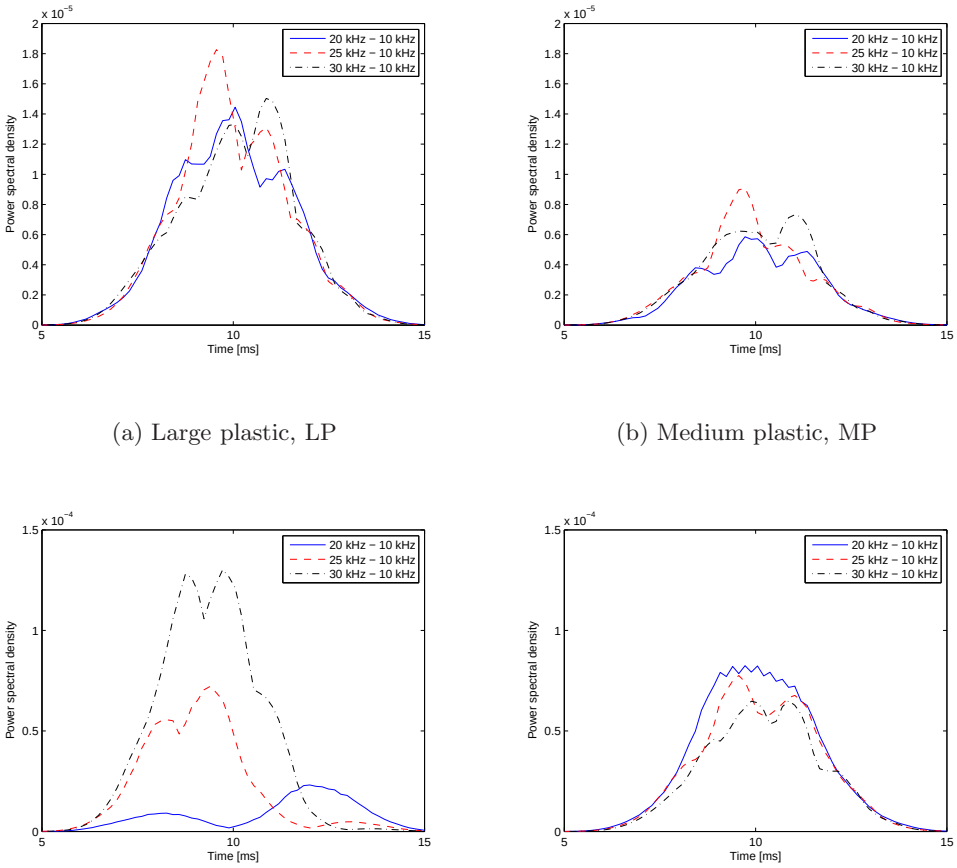


Figure 3.10: The maximum PSD of original FM chirp, 30 kHz - 10 kHz



(a) Large plastic, LP

(b) Medium plastic, MP

(c) Thick wood, TW

(d) Thin wood, Tw

Figure 3.11: The maximum PSD depending on specified sweeping ratio for LP, MP, TW and tW

To examine the centroid movement in PSD further and quantify the evaluation, the broad concept of 'activated frequency channel' has been adopted from a previous study [Boonman and Ostwald, 2007]. The 'activated frequency channel' was used in the computational model of hearing in bats. It has been defined as the number of different frequencies which stimulate the inner hair cells. In previous study, the 'activated frequency channel' was examined by investigating the peaks in the echoes which appeared in the analysis of hearing model and the number of frequencies which have a peak above a given threshold was calculated by Boonman and Ostwald [2007]. The current simulation model does not include the hearing model so the application is not exactly same but the concept of 'activated frequency channel' is adopted as an alternative tool to investigate how the simulated FM signals represent the information in the echoes. It is intended to identify the number of frequency channels being used by applying the custom-defined threshold value. The schematic description of 'number of activated frequency channels' is shown in Figure 3.12 and the procedure to define the current threshold is described later.

The analysis of a number of activated frequency channels for the LP, MP, TW and tW targets are presented in Figure 3.13. Only frequency channels with a PSD above a threshold of $2\times10^{-7}V^2/Hz$ were selected, with energy below this level counted as noise. This gives approximately a 20 dB signal-to-noise ratio for the highest amplitude level observed in the spectrogram from the different objects. As shown in the results, more frequency channels are obtained as the broader bandwidth signal is applied. In each case, a higher sweeping ratio results in more frequency bins being activated in each time segment. These effects are shown to be more obvious in the wooden discs than in the plastic discs. It also appears that, as the sweep rate is increased, the biggest increase in the number of frequency bands activated is at the shortest time frame.

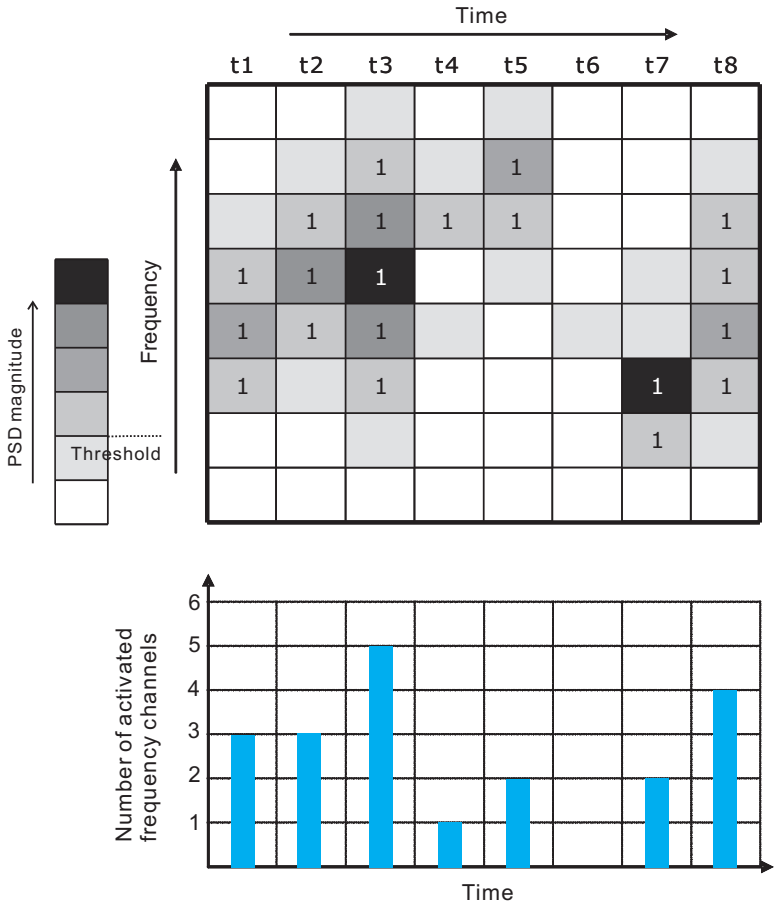


Figure 3.12: Schematic description of 'number of activated frequency channels'. This shows an example of spectrogram (top) and number of activated frequency channels against time (bottom). According to threshold defined (see colourbar on top-left), number of frequency bins which contains the PSD energy above the threshold is counted (denoted as '1').

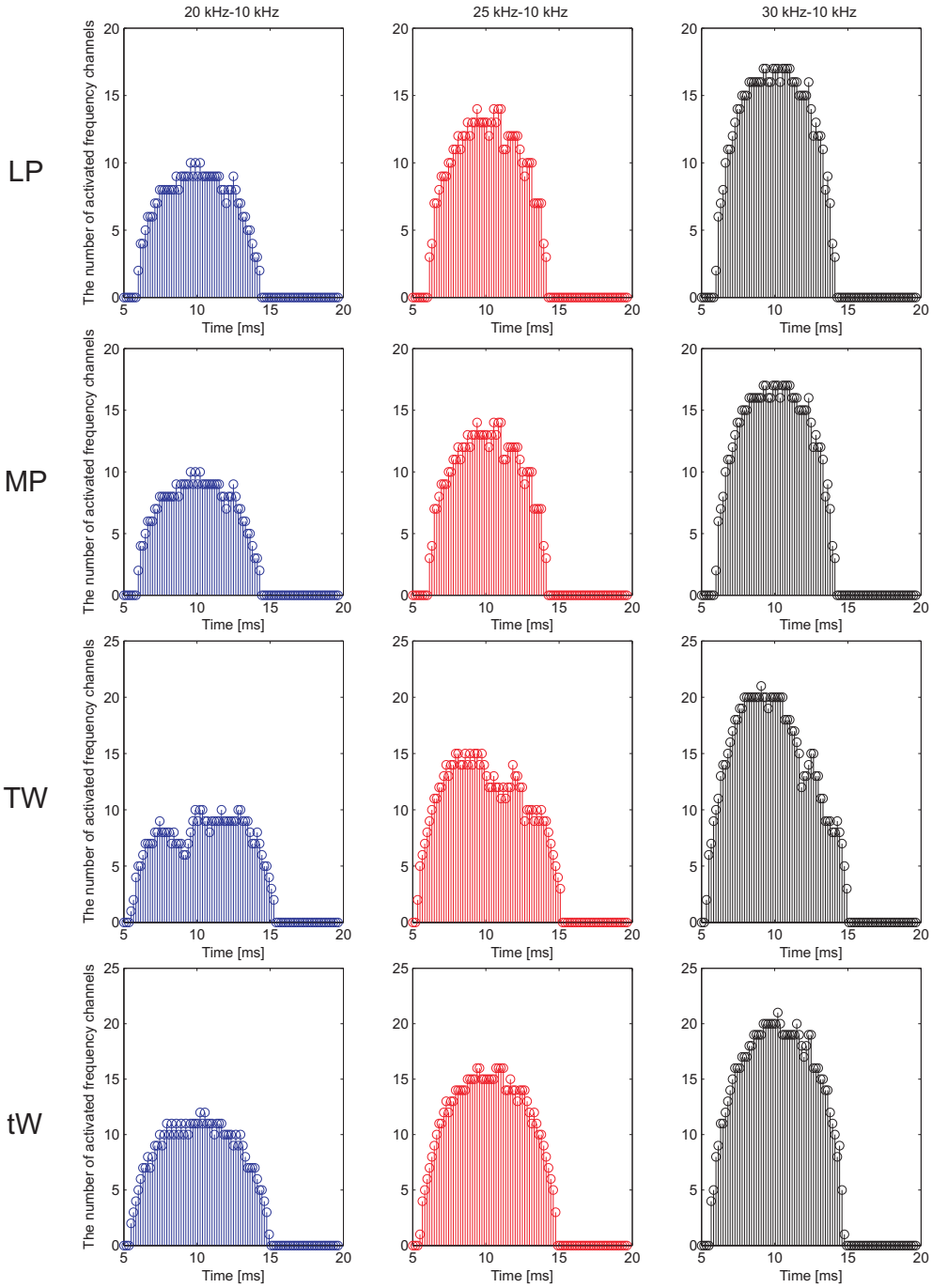


Figure 3.13: The calculated number of activated frequency channels with PSD above a threshold of $2\times10^{-7}V^2/Hz$ for LP, MP, TW and tW.

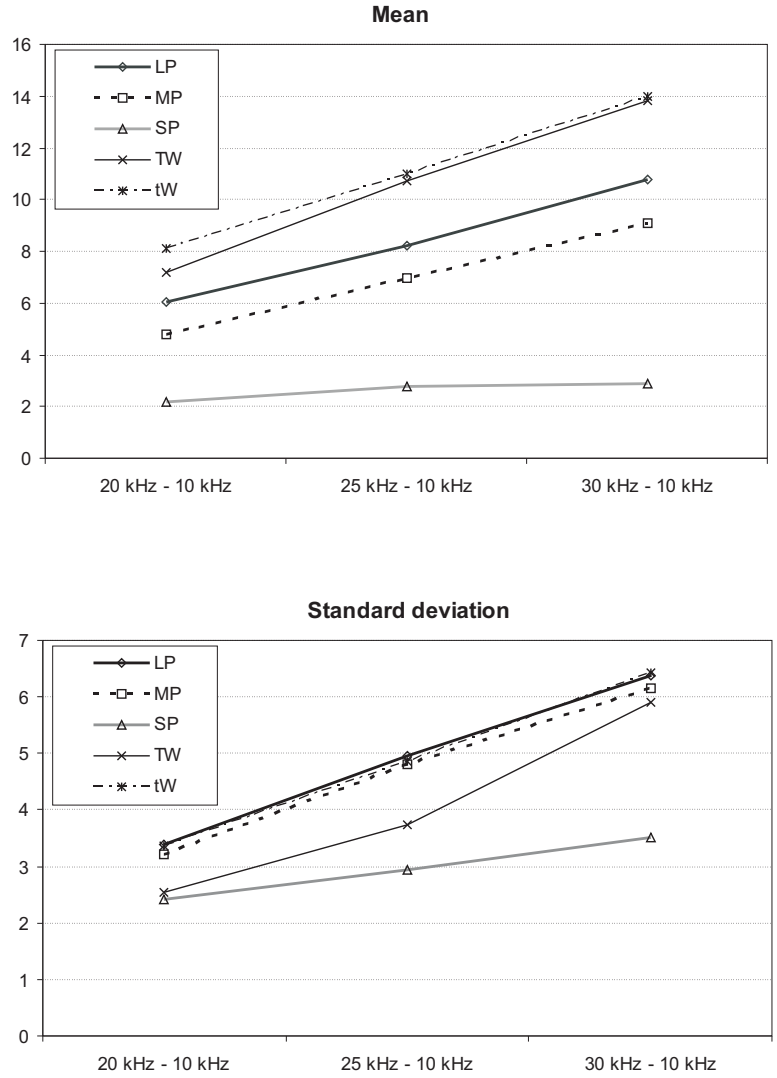


Figure 3.14: Mean and standard deviation of the number of activated frequency channels.

The mean and standard deviation of the number of activated frequency channels are evaluated for each disc as shown in Figure 3.14. Both mean and standard deviation increased as higher sweeping ratios were applied for all discs examined. Moreover, the mean increased as the diameter of the disc was increased. The result implies that the activated frequency channels provide a better explanation of the influence of the sweeping ratios than the peak characteristics and the energy distribution. However, the tW disc's mean and standard deviation are larger than those of the TW disc. It may be that the mean and standard deviation of the activated channel does not contain thickness information. However, it showed larger values for the disc with larger diameters

for all examined discs that it can be concluded that at least the thickness influences less on the result than the diameter does. Further investigation seems to be required to isolate the effects of thickness and diameter. It can be concluded that the number of different frequencies that contain energy above the noise level provide information to discriminate the size of the target although the information about the individual frequency is not provided. The larger bandwidth signal has the most activated frequency channels and thus may provide the most information on disc size.

3.1.5 Discussion

Based on the acoustic model for the echo generation mechanism, the part of the model which consists of the emitted signal and measured echo from various types of target discs has been simulated. Different signals were investigated to understand the effective structure of an echolocating signal. The spectrogram analysis of the different target discs showed different amplitude levels depending on the size of the target when the CF signal was modelled. The profile of the spectrogram shown in the FM signal provides the information on the properties of the target. The concept of the activated frequency channels being used in the differentiation of the targets for the FM signals was concluded to provide a better explanation of the spectral characteristics in the size discrimination when the FM signals were used. The results obtained using FM signals have demonstrated that a broadband signal carries better spectral information and the sweep rate affects the profile of the object differentiation pattern. It also allows more flexibility in designing FM signals in relation to bandwidths and sweeping ratio than in designing the CF signals as they are limited to the narrowband analysis, particularly in the object differentiation model. However, the result has not shown clear evidence about differentiating the thickness using either the CF signal or FM signal. But the diameter is shown to be differentiated reasonably well for both signals. To validate these results, a wider range of object dimensions and materials are required to be investigated. In conjunction with this, development of the spectrogram's pattern characterisation technique will improve target property recognition.

3.2 Hearing model

3.2.1 Introduction

The auditory system in bats is structured in the same way as that of other mammals as reviewed in Chapter 2. Once the echolocating signal returning from the target object has reached the two ears, each processes both time and frequency information in the cochlea. This study has limited the modelling to a single receptor of one input as

the binaural processing information was not available (as discussed in chapter 4). The auditory processing has been simplified based on the bandpass filter banks which model the cochlear function in the previous study by Patterson et al. [1992]. According to Patterson's model, each filter bank is modelled using a gammatone characteristic, the impulse response of which is defined in the time domain as Eq 3.2,

$$g(t) = a_g t^{(n-1)} e^{-2\pi bt} cos(2\pi f_c t + \varphi_g) \tag{3.2}$$

Where n is the filter order, b determines the bandwidth of the filter and f_c denotes the centre frequency of the filter. Also a_g determines the amplitude coefficient for gammatone impulse response and φ_g determines the phase delay. The equivalent rectangular bandwidth (ERB) is a psychoacoustic measure of the width of the auditory filter at each point along the cochlea, and can be defined as Eq 3.3 [Slaney, 1993];

$$ERB = \frac{f_c}{EarQ} + minBW \tag{3.3}$$

where $EarQ$ is the asymptotic[1] filter quality. When the frequency reaches the high frequency region, $minBW$ can be neglected. In contrast, $minBW$ is the minimum bandwidth possible, and occurs when $\frac{f_c}{EarQ}$ approaches 0. Thus, the $minBW$ term can be neglected as f_c becomes larger. For human auditory modelling, Glasberg and Moore [1990] have recommended data on the equivalent rectangular bandwidth of the auditory filter for the Eq 3.3 with $EarQ = 9.26$ and $minBW = 24.7$. However, there is no direct evidence of $EarQ$ values for bat's auditory modelling. Assuming the auditory processing in bats resembles that of humans in terms of their mammalian cochlea, these values has been adopted and modified for the auditory filter banks of both FM and CF bats. This provides a simple and bio-inspired auditory model based on the more complex bat's system.

3.2.2 Auditory model

FM bats: general model

FM bats are thought to possess a similar general pattern of auditory filter banks as other mammals. Therefore, the 'general model' based on the human cochlear model is designed to process the FM signals. Assuming better performance in resolving frequencies in the cochlear of the FM bats than in those of humans, the $EarQ$ factor is varied as 10, 20 and 30. The higher $EarQ$ represents the sharper tuning curve for each frequency in the auditory system. The $minBW$ is set to the same value as in Glasberg

[1] $\frac{f_c}{EarQ}$ changes from almost negligible to a very large number compared to $minBW$ as f_c gets larger.

and Moore's equation as it is negligible in the frequency range of the analysis [Glasberg and Moore, 1990]. The modified frequency response of the gammatone filter bank for an $EarQ$ value of 10 is shown in Figure 3.15(a). The equivalent rectangular bandwidths used for the cochlea modelling of the FM bat with different $EarQ$ values are shown in Figure 3.15(b). The bandwidth becomes more narrowly tuned as the $EarQ$ values increases.

The gammatone filter banks were modelled within the frequency range from 5 kHz to 35 kHz as the high frequency limitation of the current system is 40 kHz. Each bandpass filter bank produces the demodulation of the frequency seen through the basilar membrane. The output of a linear filter bank resolves the frequencies to some extent, yet it does not represent the neural activity in the cochlea. After the frequency is resolved by the mechanical motion in the basilar membrane, the inner hair cells convert the motion into neural activity. In the first instance, extraction of the signal envelope from the motion by the inner hair cells is modelled by half-wave rectification and low pass filtering. As the rectifier generated new frequency components, some of which are unwanted by the specific frequency channel, low pass filters at 1000 Hz were chosen to remove these unwanted components. The output is extracted from the envelope of the decomposed signal. The envelope of each channel is applied to the square root compression which reduces the fluctuation of the enveolpe. The square root compression processing represents the gain control system in the cochlea. As a result, the output signal is fitted within the dynamic range of hearing.

CF bats: Acoustic fovea model

CF bats are thought to carry out specialised frequency analysis in the acoustic fovea for Doppler shift compensation. In this region, the neurons are specifically tuned to a narrow frequency range of the returning echoes which are modified from the emitted signal. In this study, the narrowband filters are designed separately to model the acoustic fovea where the analysis of the CF signal is performed. Therefore we aim to investigate how advantageous this mechanism is in terms of object discrimination. For both cases (with/without the foveated model), the bandwidth b is fixed as 1.019 times the ERB, as recommended by a previous study [Slaney, 1993].

The narrowband filters are designed to reconstruct the processing in the acoustic fovea of the cochlear of CF bats. For the 20 kHz carrier frequency of the echolocating signal, the acoustic fovea has been assumed to be located from 18 kHz to 22 kHz based on data from *Pteronotus parnellii* as described in Chapter 2. The $EarQ$ was specifically

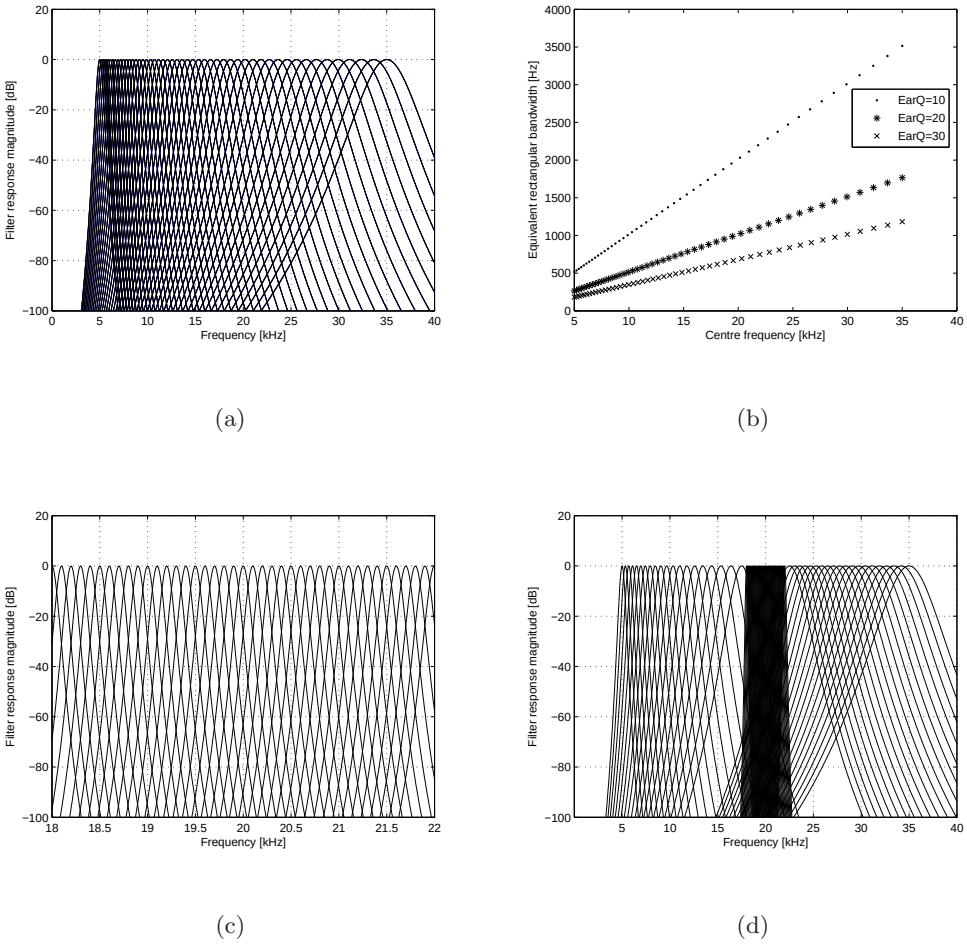


Figure 3.15: (a) Modified gammatone filterbanks of auditory processing in FM bats based on Glasberg and Moore parameters (1990) (b) Equivalent rectangular bandwidths (ERB) with different $EarQ$ values (c) The narrowband filterbanks used to model the acoustic fovea of CF bat (d) The modified gammatone filterbanks of auditory processing in CF bats

changed for the narrowband analysis. A fixed $EarQ$ value of 100 was chosen which is approximately ten times larger than that of the human model. The separation between each centre frequency of the narrowband region was set as 100 Hz. The filters were also spaced to give equal overlapping between the neighbouring filters. Hence the acoustic fovea model generates 41 narrowband filters and their frequency responses are shown in Figure 3.15(c). Outside the acoustic fovea, including adjacent frequencies, the frequency analysis was divided into 5 kHz-17.5 kHz and 22.5 kHz-35 kHz. Each part was modelled based on general gammatone filterbanks with the $EarQ$ value of 10. The filter response is whole frequency region is shown in Figure 3.15(d).

3.2.3 Reconstruction and analysis of echolocation signals

The signal processing was performed using MATLAB version 7.2 and the flow diagram of the processing is shown in Figure 3.16. Here, the reconstruction and signal analysis procedure consists of two: simulation of the echolocating signal in the air transmission and auditory pattern spectrogram-like analysis. The first step is to reconstruct FM and CF echolocating signals. The FM signals were generated with the bandwidth which was set to sweep down from 30 kHz to 10 kHz in 10 ms durations and the CF signals of 20 kHz sinusoidal waves in 10 ms duration were generated. Each signal was multiplied by a Hanning window, which has same length as the signal, to remove the undesirable high frequency components at the beginning and end of the signal. The generated signals were convolved with each backscattering impulse response from the discs and auditory processing was performed for each signals. Finally, the auditory pattern spectrogram-like image was processed for each signal's auditory model and plotted as a contour image.

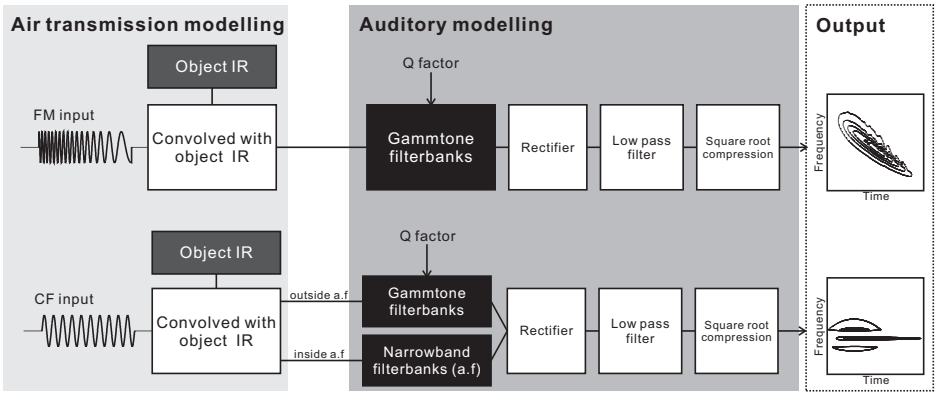


Figure 3.16: Diagram of auditory processing. 'a.f' denotes acoustic fovea.

Then, the image subtraction of the spectrograms from two different objects is used to show the auditory pattern characteristics for object differentiation [Kim et al., 2008]. The difference between the outputs of the auditory processing for two objects was calculated and presented. The difference in auditory pattern contour enables us to identify the characteristics of object differences more easily than the individual disc auditory images.

3.2.4 Convolved signals

Figure 3.17 shows the impulse response from each disc, the convolved outputs using FM and CF signals, and their envelopes. The impulse response has shown the difference in peak level depending on the dimension of the discs. It is obvious that the disc which has

larger diameter produces higher amplitude. But it has also shown that the thicker disc (TW) produces higher amplitude than the thinner disc (tW). The convolved outputs simulates the signals of the received echo from each disc and they resembles the shape of hanning window for both FM and CF signals according to the signal processing carried out in this study. The FM and CF signal generates different shape of the envelopes in response to the frequencies contained in each signal. The FM signals show various peaks and notches in the envelopes whereas the CF signals show a single peak for each disc. For both cases, the energy contained in the envelope is calculated and plotted over the diameter of the discs and the result is shown in Figure 3.18. The FM signals show around 25 dB of dynamic range for the given range whereas the CF signals show around 30 dB of dynamic range. The difference of magnitude between TW and tW is larger when CF signal is used than when FM signal is used. The result implies that the relationship between the diameter and the envelope magnitude of object's echo is stronger when FM signals are used, hence the effect of the thickness is suppressed.

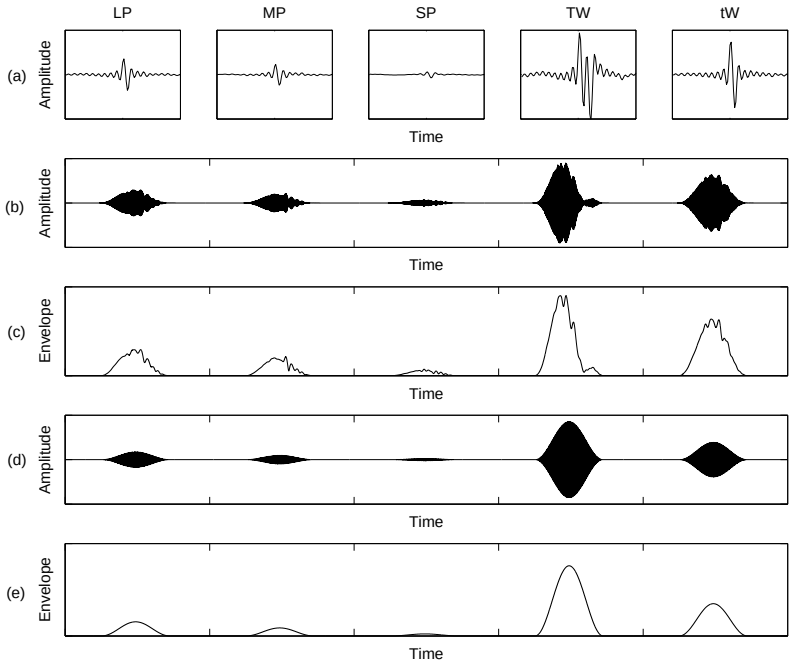


Figure 3.17: (a) Impulse responses from 5 different discs (b) convolved signals for FM (c) envelope of convolved signals for FM (d) convolved signals for CF (e) envelope of the convolved signals for CF

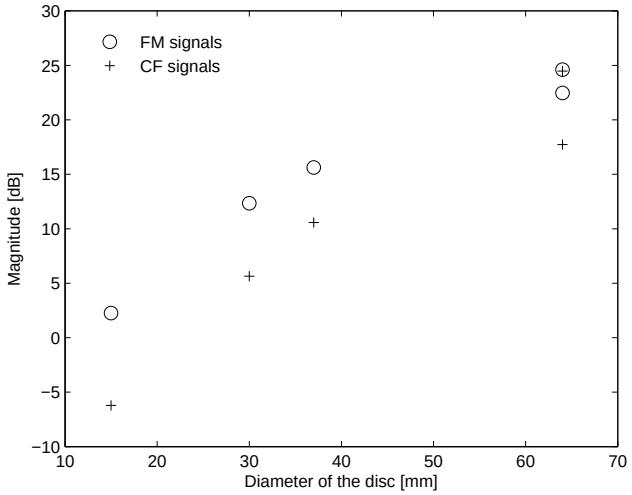


Figure 3.18: Comparison between FM and CF signals for the energy contained in the envelopes

3.2.5 Results

FM signals vs. CF signals

The processed auditory images are presented for both FM and CF signals. Figure 3.19 plots the auditory image contours for 5 different discs when FM signal is applied. As expected, the contours show that the diameter of each disc can be discriminated by the intensity. The larger object produces the higher amplitude in the reflected signal. Although the discrimination between each object is available with $EarQ$ factors of 10, 20 and 30, the contour image is more sharpened as the $EarQ$ increases. It is considered to be due to the fact that the energy in each frequency component produces less leakage into the adjacent frequencies. On the other hand, the envelope analysis based on the each contour image is conducted by summing up the energy in each time segment, then similar shape of the signal envelope shown in time domain is obtained. However, as the $EarQ$ increases, less amplitude level of envelope was produced due to the narrower band analysis from higher $EarQ$ which induces low level of filtered energy. It is also noted that the envelope converges into a single peak as the $EarQ$ increases. Also the difference between $EarQ$ values of 10 and 20 is larger than that between $EarQ$ values of 20 and 30.

Figure 3.20 shows the results from the processing of CF signals. Similarly, diameter of the disc can be discriminated by the intensity shown in the auditory image contour. For the CF signals, the effect of acoustic fovea model is investigated. It is shown in the contours that 20 kHz component of CF signal is well separated from the adjacent

frequencies when the acoustic fovea model is applied. The foveated model produced higher resolution of analysis in the 20 kHz frequency region, hence the auditory image was more sharpened compared to the results from the general model. It is interesting to note that the peak of the envelope is delayed when acoustic fovea model is applied. It is considered to be due to the fact that the foveated model possesses larger number of filterbanks compared to the general model in the same frequency region. Therefore, it is expected that the signal which is processed through larger number of narrowband filters, will show relatively more energy because of their overlapping effects between adjacent frequencies.

It is also interesting to compare the performance of object discrimination between the FM and the CF signals. First of all, both results show similar bell-shaped patterns in their envelopes although the contour of each signal appear different. However, the envelope of the CF contour has appeared to contain less noise than that of FM contour when same $EarQ$ value of 10 is applied.

The envelope magnitude is extracted along the time axis and its level is calculated in decibels by summing the sound energy in each time segment. Figure 3.21 demonstrates the magnitudes of auditory envelopes calculated from the results of the FM signals for $EarQ$ factors of 10, 20 and 30, and the CF signals for general model foveated model. The results plots over the diameter of the dics. The dynamic range are shown to be approximately 10 dB for all cases although there are level differences depending on the type of signal and processing method. All results show that the magnitude increases proportionally as the diameter of the disc increases. For CF signals, it has been shown that the foveated model appears to provide larger magnitude by approximately 3 dB than the general model. The result implies that the foveated model prevents the leakage of auditory energy in the CF echolocating signal. For FM signals, the higher $EarQ$ value produces the lower level of magnitude. It is considered to be due to the fact that the bandwidths of gammatone filterbanks are narrower when the higher $EarQ$ is applied. It is interesting to note that the magnitude gain of CF signal is generally higher than that of FM signals, however the result from the FM signal of $EarQ$ factor of 10 was almost same as the result from CF signal in general model. This implies that same performance of auditory processing can be achieved although different types of signals are used. Lastly, the magnitude difference shown for the two discs in different thickness but same diameter of 64 mm, were prominently suppressed when the auditory processing is applied. Note that this result is comparable to Figure 3.18 which has shown the envelopes of the raw signals in time-domain.

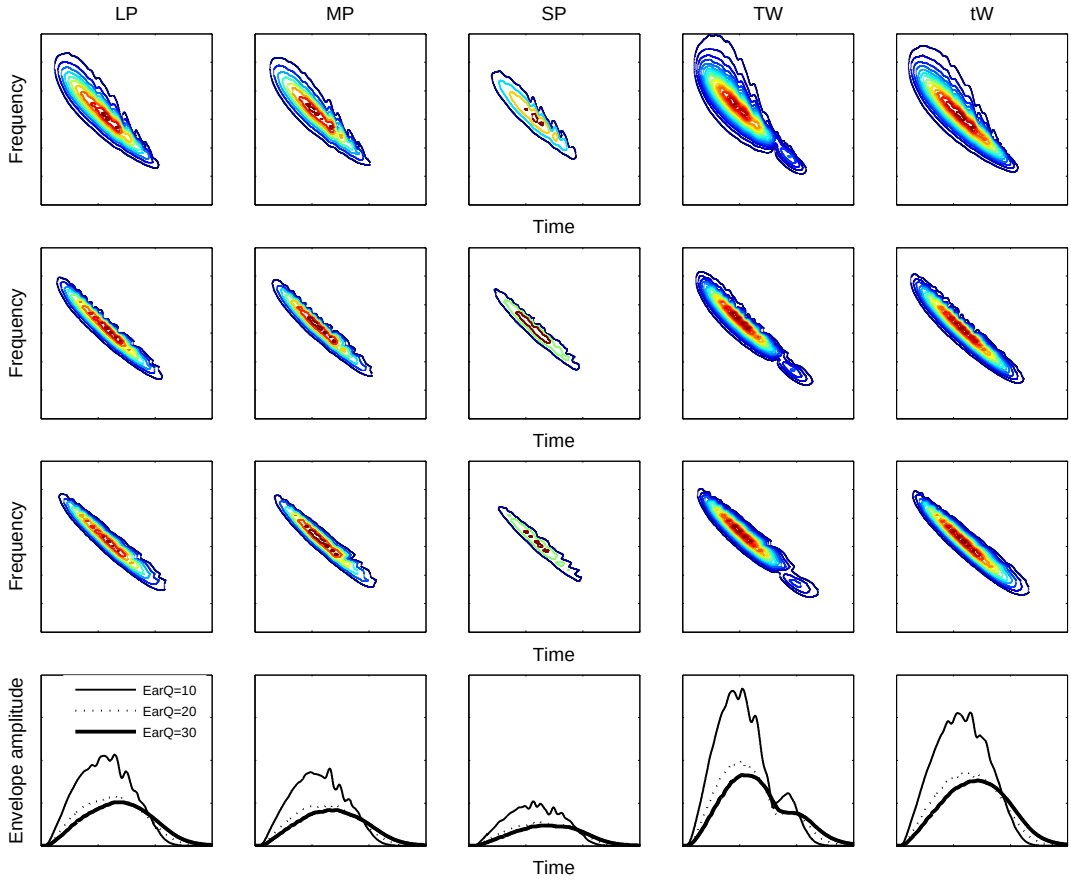


Figure 3.19: Auditory image of 5 different discs are plotted when FM signal are applied. First row: $EarQ = 10$, second row: $EarQ = 20$ and third row: $EarQ = 30$. The last row shows the envelope magnitude of given auditory images in different $EarQ$ factor for each disc. The envelope magnitude was calculated by summing up the energy contained in each time segment over the frequency.

Object discrimination - image subtraction

The subtracted auditory pattern images for the LP-SP discs and the TW-tW discs are presented for both FM and CF signals when different types of auditory processing were applied. Figure 3.22 shows the pattern how the difference between two object can be demonstrated depending on type of signal and the auditory processing method. For FM signals, a higher $EarQ$ results in frequency representation of higher resolution in each time segment, therefore, it sharpens the auditory image. For the CF signals, the foveated model enables to separate the echolocating frequency component in the contour. However, it is noted that the results from TW-tW discs using CF signal were shown to have no prominent feature for both cases of the general and the foveated

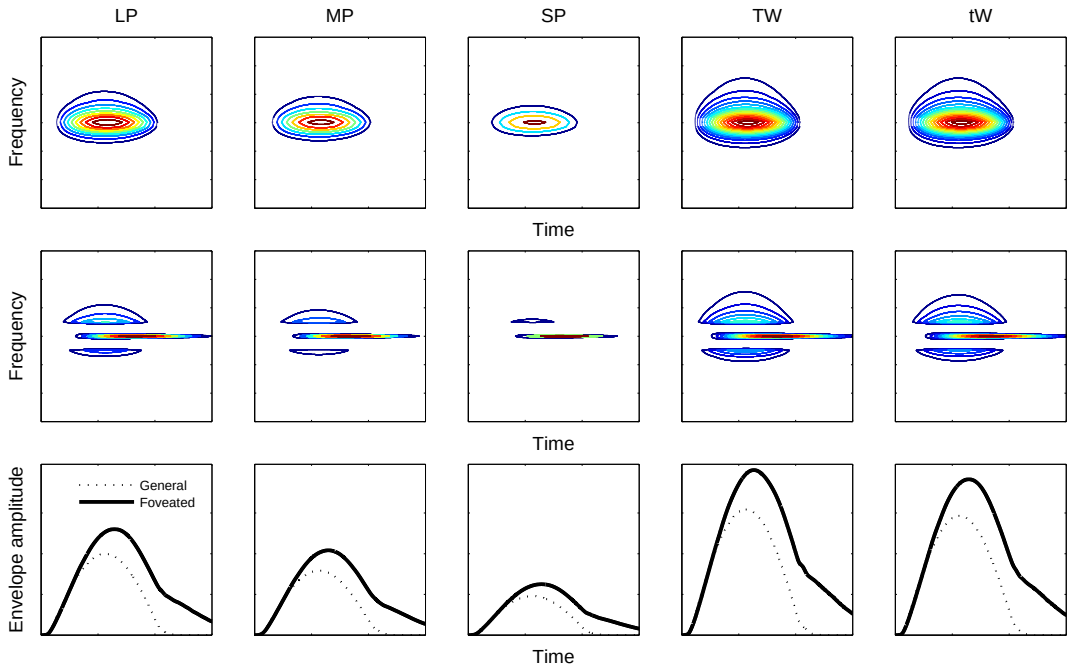


Figure 3.20: Auditory image of 5 different discs are plotted when CF signal are applied. First row represents the result from the processing without acoustic fovea and second row represents the result from the processing with acoustic fovea. The last row shows the envelope magnitude of given auditory images. The envelope magnitude was calculated by summing up the energy contained in each time segment over the frequency.

model, whereas the results of FM signal clearly produced a pattern of differentiation in the contours. This indicates the better performance of the FM signals in terms of thickness discrimination compared to the CF signals.

3.2.6 Discussion

Based on the acoustic model of the echo generation mechanism, we have simulated the bat-oriented auditory processing of echolocating signals in terms of the object differentiation. The bat-oriented gammatone auditory filterbanks were designed based on the standard mammalian cochlear model and modified for both FM and CF signals. As a result the difference between the outputs from the cochlear processing of two objects produced a more sharpened image for the higher $EarQ$ value in the processing of FM signals and it resulted in convergance of the envelope into a single peak. The sharpened image indicates the necesity to process narrower frequency bandwidth for each time segment which might reduce the leakage of frequency energy into adjacent time segments. Higher $EarQ$ also limit the distortion in the envelope and enhance the

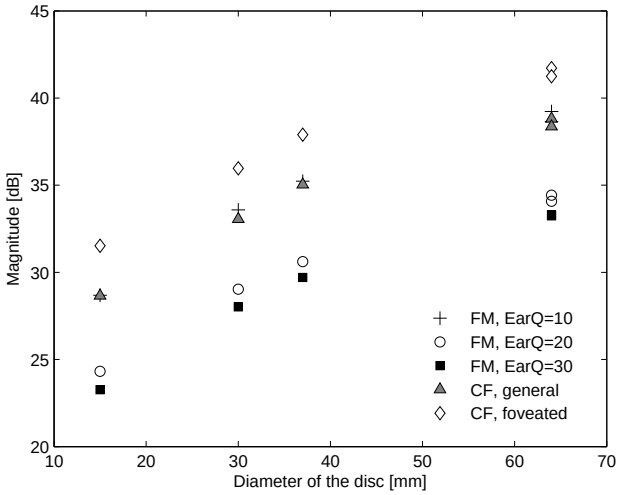


Figure 3.21: The magnitude of the envelopes from auditory processing. The performance of each processing with FM/CF signals are compared by calculating the energy of each envelope.

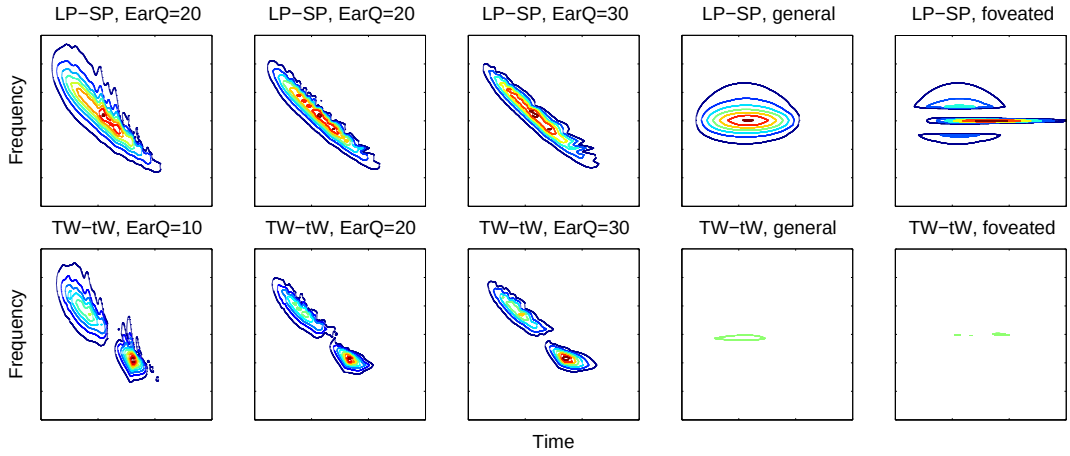


Figure 3.22: LP-SP discs difference and TW-tW discs difference in the auditory pattern spectrogram-like image

spectral contrast. On the other hand, the auditory image from the analysis of the CF signal has presented similar characteristics in the spectra compared to the processing of FM signal, in terms of envelope representation. But the foveated model has produced higher frequency resolution and provided a sharpened image. Moreover, the foveated model has enhanced the sound energy in the signal by approximately 3 dB compared to the general model. In summary, the results imply that the FM signal is more suitable for object characterisation than the CF signal. Yet, the processing in the fovea is advantageous in terms of increasing the auditory amplitude level.

In terms of disc discrimination, the current auditory processing enhanced the performance in discriminating the diameter of the disc. This study suggests that both auditory processing in FM and CF signals compress the input dynamic range of the envelope magnitude in the returning echoes. The comparison of the envelope magnitude between the signals before and after the auditory processing show that the dynamic range was compressed from 25-30 dB to 10 dB following the processing. Moreover, the auditory processing suppressed the effect of thickness on the envelope magnitudes so that the temporal analysis appeared more susceptible to changes in the diameter of the disc. The discs of different diameters showed the proportional increase in envelope amplitudes when the CF signal was applied. Using the temporal analysis, both FM and CF signals were able to provide enough information to discriminate the overall size, but information about the particular shape, thickness or material was not clearly observed. This study has shown that the thickness information is carried by subtracting the auditory image of the two different discs. However, this effect was limited to the FM signals and the thickness difference was not clearly found in the auditory image of CF signals. This result suggests that diameter is determined by temporal processing, whereas the thickness information might be encoded in the spectral features of echoes
.

3.3 General discussion

The simulation results demonstrated how different types of echolocating signals can be used to discriminate discs. In this study, data was available to discriminate diameter but it was difficult to find obvious evidence about how to discriminate the thickness and material from both CF and FM signals in the spectral analysis. The discs of different diameters show the proportional increase in amplitude with the simulated CF echolocating signal. The CF signal could provide enough information to discriminate the overall size but it does not seem to carry information about the particular shape, thickness or material.

The temporal information was also observed in the convolved signal for each object. As demonstrated earlier, the convolved signals of CF showed generally similar envelope patterns between all discs tested and no prominent difference was found except the magnitude difference. Therefore, the temporal analysis supports the hypothesis that a CF signal of a single frequency is limited to discrimination of geometrical size when a target is only viewed from the front. In contrast, it is notable that the pattern of the

convolved FM signal appeared in the different patterns for each target disc, at later time periods especially. It can be assumed that size-independent characteristics such as shape, thickness and material are also encoded in the FM signal. Simmons and Chen [1989] have shown that the structure of the impulse responses from mealworms, discs and spheres provide the profile of discrimination to FM echolocating bats. Simon et al. [2006] have shown the systematic size-dependent changes in the echo's temporal and spectral pattern as well as in the amplitude by looking at the impulse response of different sizes of a hollow hemisphere. This study also supports the theory that the temporal patterns might provide a better explanation for the target profile. However, the current study gives great emphasis to the temporal pattern of the type of signal being used. In other words, the encoding of characteristics in the impulse response from a particular object can be varied depending on the type of the signal used. Moreover, this study has shown that the bandwidth of the FM signal has appeared to affect the characteristics of the target profile. Broadening the bandwidth in the fixed duration improves the number of activated frequency channels in the earlier timing stage of the echo. This implies that the information concerning the object's characteristics contained in a single echo may avoid overlapping with that of following echo when a bat emits multiple echolocating pulses.

The auditory processing in this study contributes to enhancing the performance of each signal. In general, the higher gain is achieved and the dynamic range is compressed when the auditory model is applied to the simulated echo signals. In addition, the auditory processing with higher quality filters improves the resolution of the temporal pattern when an FM signal is used. Also it contributes to an increase in the amplitude level in the auditory system when the CF signal is processed through the acoustic fovea model. The result supports the role of each signal: the FM signal is advantageous for temporal processing; and CF signal is advantageous for fovea processing. On the other hand, it should be noted that there is a flexibility in designing an auditory model not only by varying the filter quality but also by varying the number of filters in a given frequency range. This variation may influence practical issues, such as computing efficiency, for example more filters in auditory processing may require more computing time depending on the performance of hardware.

Chapter 4

HRTF Measurement of a Bat-head Cast

4.1 Introduction

The head-related transfer function (HRTF) introduces the binaural information which provides the source for the localisation study. It describes the acoustical influence of the body, head and external ears. The effects of the external ears can be considered as a linear time-invariant (LTI) system. Information of the system is encoded in impulse response (in the time domain) or transfer function (in the frequency domain). The transfer function at each ear, measured at various angles and distances from the sound source, contains the information available to the subject receives in order to determine the location of the source. The human HRTF and human sound localisation has been intensively studied. The various measurement and analysis techniques being applied to determine the human HRTF provide a good reference in terms of methodology in the measurement of the HRTF in a bat head. On the other hand, there have been relatively few studies in the HRTF measurements of animals. This study is especially interested in those which have a small head size of only a few centimetres or less in diameter, including various types of bat as well as rat and gerbil. These animals often use an auditory frequency range higher than humans (above 20 kHz), therefore the issues of small head size and high-frequency measurements are considered to be important to produce reasonably accurate and reliable HRTF measurement data for such animals.

This chapter reviews the small-head HRTF measurements which have been conducted in previous studies Jen and Chen [1988]; Obrist et al. [1993]; Fuzessery [1996]; Aytekin et al. [2004]; Firzlaff and Schuller [2004]; Maki and Furukawa [2005]; Koka et al. [2008] and examines the current measurement system. The repeatability of measurements in a

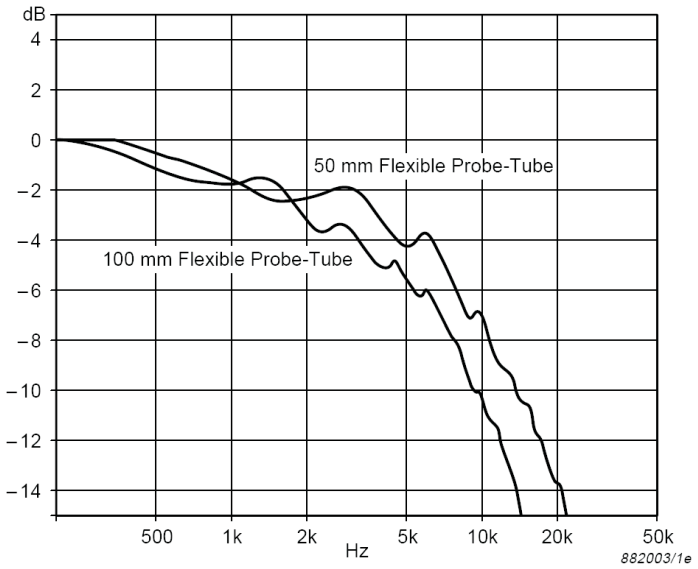
bat-head cast is investigated to validate the measured data. In section 4.2, the previous studies of small-head animal HRTF measurement are reviewed in terms of reliability of the system. In section 4.3, the measurement methods and the current system response are examined. In section 4.4, the HRTF measurement using the bat-head cast is conducted and the repeatability is assessed. The general characteristics of the HRTF measured in the bat-head cast are also shown and discussed.

4.2 Small-animal HRTF measurement

There are couple of issues to be discussed to produce the reliability in the HRTFs measurement using small-headed animals. Firstly, it is necessary to carefully consider the method of fixing the microphone appropriately to the ear position of the head to measure the HRTFs. The microphone positioning within such a small head (a few centimetre in diameter) is not easy because the physical volume of the microphone is relatively large to be positioned inside the head. For example, the human head model, KEMAR [1], accommodates two microphones lying down at the position of each ear drum but a 4 cm diameter bat-head cast is difficult to accommodate even just one microphone. As a result, most of the previous studies have used some kind of extension tube which is attached to the microphone. As shown in Table 4.1, the extension tube varies from custom-made plastic, silicon or metal, to the probe microphone manufactured professionally. It is commonly agreed that the frequency response of the microphone with an extension tube has to be examined and characterised before the measurements are taken. A probe microphone from industry seems to be an easy solution from which to build the extension and measure the response. Yet, it generally does not support the smooth and flat frequency response above 20 kHz, and the response drops off quickly above a particular frequency. An example of the frequency response of a manufactured probe microphone by B&K is shown in Figure 4.1. The product data only shows the frequency range up to 20 kHz and its response is shown to have low magnitude under -15 dB and is expected to be even lower after 20 kHz. The B&K probe microphone has recently been used in previous studies to measure the HRTFs of the gerbil and the rat [Maki and Furukawa, 2005; Koka et al., 2008]. It is considered necessary to examine the response above 20 kHz as both studies have allowed the frequency analysis up to 50 kHz and 40 kHz each. Although it has been possible to obtain the response up to

[1]KEMAR Manikin is head and torso simulators (HATS) based on worldwide average human male and female head and torso dimensions and meet the requirements of ANSI S3.36/ASA58-1985. The manikins simulate the changes that occur to sound waves as they pass a human head and torso such as the diffraction and reflection around each ear (http://www.gras.dk/00012/00330/)

those frequencies in previous studies, this particular microphone does not seem to be suitable for the current bat-head as it is unable to cover the frequency range of HRTF measurement which requires up to 100 kHz or even more.



(a) B&K probe microphone 4182, frequency response

(b) B&K probe microphone 4182, photo

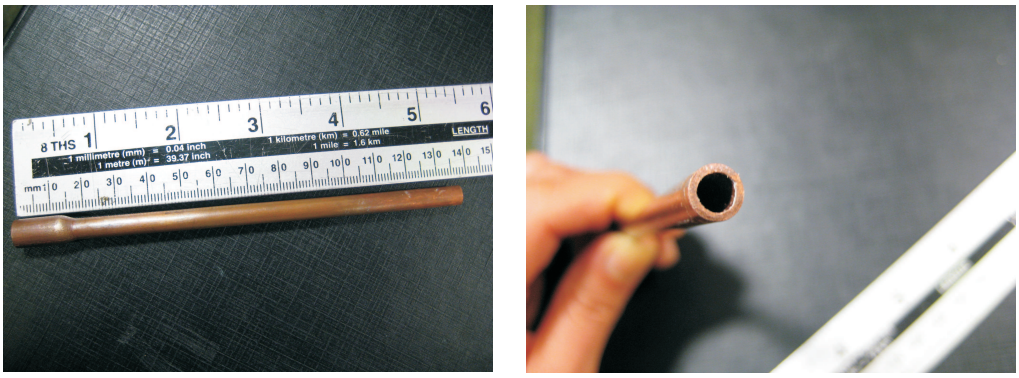
Figure 4.1: Manufactured probe microphone B&K 4182, taken from B&K product data (a) frequency response (b) photo of the probe

Therefore, the custom-made tube extension probably has been widely used in other studies for the measurement of the bat-head [Jen and Chen, 1988; Obrist et al., 1993; Aytekin et al., 2004; Firzlaff and Schuller, 2004]. However, the tube extension has a high chance of introducing resonant peaks and distortion as the acoustics in the tube generates reflections inside when the sound is transmitted. A few previous studies have

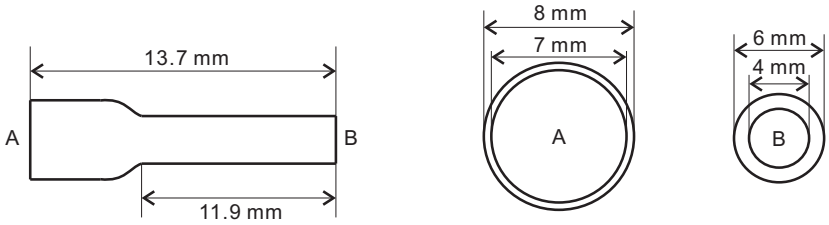
shown that these effects can be cancelled out by post-processing, e.g. by free-field equalisation or calibration. Free-field equalisation is the process of eliminating the effect of the microphone and the speaker and it can be conducted by subtracting the frequency response of the microphone and the speaker from the measured response. However, in those studies, the original system capability has not been described in detail to show the reliability for obtaining raw data of good quality in the first place. This is considered to be important in terms of signal-to-noise ratio (SNR) due to the fact that the ILD can be as large as 40 dB or more [Erulkar, 1972; Harnischfeger et al., 1985; Obrist et al., 1993]. For example, if the magnitude of the frequency response rolls off quickly beyond high frequency region, it is difficult to use the ILD data from those frequencies with reliability. It means that the raw data can be distorted when the SNR is not high enough in the frequency region of analysis. Therefore, the free field equalisation does not work effectively any more because of high noise level contained in the data.

In this study, the effect of a custom-made metal tube is investigated. The metal coupler has been designed and used in previous studies to measure the HRTF of the spear-nosed bat (*Phyllostomus discolour*), and CF/FM bats (*Pteronotus parnellii* and *Rhinolophus rouxi*) [Firzlaff and Schuller, 2003, 2004]. We designed a modified tube as shown in Figure 4.2. The top end, which fits into the ear position of the bat-head cast, has an inside diameter of 4 mm, and an outside diameter of 6 mm. The bottom end, which is fixed to the microphone, has an inside diameter of 7 mm, and outside diameter of 8 mm in the end. It is made of copper and inserted into the bat-head cast to receive the sound. The straight section 11.9 cm long extends to the position of the entrance to ear canal of the bat cast. The length of the straight section (11.9 cm) provides a distance which is long enough to attach the microphone outside of the bat cast. The acoustical effect of the coupler can be compensated by the measurement and equalisation of the free-field response as described above. This is discussed in section 4.3.

Secondly, the distance between the head and the speaker varies between most of the studies as shown in Table 4.1. In human studies, the conventional 1 m reference of distance between sound source and subject has been used. In a previous study [Brungart et al., 1996], the 1 m reference is theoretically supported by the modelling study which has shown that the HRTF changes substantially as the distance varies at less than 1m, yet are independent of distance of source beyond 1 m. Therefore the terminologies 'proximal region' for the region less than 1m and 'distal region' beyond 1 m has been suggested to describe the distance effect on the listener's head [Brungart and Rabinowitz, 1999]. On the other hand, the distance effect seems to be more important



(a) (b)



(c)

Figure 4.2: Custom-made metal tube and its dimensions (a) photo in whole length (b) photo from the top end (c) dimensions, A: bottom-end (connected to the microphone) B: top-end (where the sound is recieved)

particularly for the investigation of the HRTFs in bats. Because they always fly towards the prey which means that they utilise different HRTFs at various distances for echolocating behaviour. However, there has not been a consideration of the reference distance for the HRTFs measurements. Obrist et al. [1993] used a 15 cm distance whereas Koka et al. [2008] used 1 m distance and other studies have fixed the distances between them. Based on the human study described above, the normalised distance can be adopted and it shows the expected referencing distance according to the different sizes of head. The normalised distance (ρ) is defined by Eq 4.1 as a function of distance from the centre of the head to the source (r) and radius of the head (a).

$$\rho = \frac{r}{a} \tag{4.1}$$

Then the normalised distance is set as 10 for the human study with the assumption that the reference distance is 1 m and the radius of the head is 10 cm. The reference distances for the various sizes of the smaller heads are then calculated using the normalised distance value of 10. The area of proximal region and distal region is shown

in Figure 4.3. It means that the expected proximal region is about 20 cm for the 2 cm radius of head size. It can be concluded that most of the small-headed animal measurements have been conducted in the distal region in previous studies, yet there seems to be no particular consideration of the proximal or distal region.

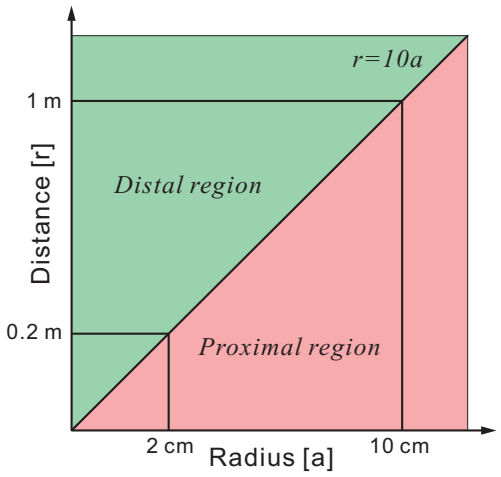


Figure 4.3: The schematic diagram of the proximal region and distal region. The distance of the proximal region reduces as the radius of the head size reduces. It shows the boundary between the proximal and distal region at 1 m for the 10 cm radius and 20 cm for the 2 cm of radius in the head size.

Thirdly, the signal type and the processing for analysis are important conditions for the HRTF measurement. There are various types of signals being used in the measurements in previous studies. In the measurements of bats, Jen and Chen [1988], Obrist et al. [1993] and Fuzessery [1996] have used pure tones of different frequencies, and Aytekin et al. [2004] has used linear FM chirps for the measurement. These studies revealed the relationship between the measured HRTFs and a particular signal which a specific bat uses. Those different signals are also used to see whether any special feature of the pinna is associated with the signal the bats use. It is obvious that there exists a limitation to observe the response to wide frequency range when a signal of limited bandwidth is used. Therefore, broadband signals such as pink noise, random noise or MLS (Maximum Length Sequence) have been widely used in human studies to obtain the frequency response. As this study is intended to investigate the general features in the HRTF of an artificial bat head rather than to investigate a specific type of bat's HRTF, the broadband signal is considered to be most appropriate.

Based on the review described above, the system requirement for the HRTF measurement of the bat-head cast is specified and it is shown in Table 4.2. Major key

specifications are as follows: distance, signal type, SNR and head positioning. Considering the head size of the bat-head cast, distance between the loudspeaker and the head is set to 90 cm which indicates the receiver is located in distal region. Broadband analysis is available by using pink noise and the SNR is set to 50 dB to support the dynamic range of interaural level differences. Also, two laser beams are projected to position the centre of the head.

The data processing methods depend on the type of signal used in the measurement. As shown in Table 4.3, the different methods used in previous studies are listed in terms of processing the ILDs and ITDs. In the analysis of the ILDs, various methods such as sound pressure level measurement for each frequency, bandpass filtering and peak-to peak gain analysis have been used when the testing signal is of limited bandwidth. On the other hand, the studies which used the broadband signal can perform a FFT (Fast Fourier Transform). The size of the FFT associates with the frequency resolution to be achieved. Windowing is also applied in advance to produce a smooth response curve. The decision to choose the processing parameters of the FFT and the type of window can be different in each study to produce better presentation of the results. In the analysis of the ITDs, only a few studies have performed the data processing and envelope cross-correlation techniques [Aytekin et al., 2004; Koka et al., 2008]. As the range of the frequency analysis is mostly over 10 kHz and the ITDs are known to be unavailable in those frequencies, extracting the envelope of the signals received at the two ears seems to be a reasonable approach for the processing. In this study, FFT method for ILD analysis is recommended as it provides straightforward results for various frequencies when broadband input signal is applied. 8192-point FFT is particularly chosen to resolve the provided frequency analysis range reasonably well. For ITD analysis, the envelope extraction method is recommended as discussed above and Hilbert transform is used for this study.

Table 4.1: Small-animal HRTF measurement reported in the literature

	Jen and Chen [1988]	Obrist et al. [1993]	Fuzessery [1996]	Aytekin et al. [2004]	Firzlaff and Schuller [2004]	Maki and Furukawa [2005]	Koka et al. [2008]
Subject	Bat	Bat	Bat	Bat	Bat	Gerbil	Rat
Head diameter	10-20 mm	10-20 mm	10-20 mm	10-20 mm	10-20 mm	17-19 mm	28-31 mm
Position of the receiver at the ear	close to the ear drum	where the ear drum locates	inner position in the ear canal	where the ear drum locates	close to the ear drum	inner position in the ear canal	entrance to the ear canal
Distance	Not known	15 cm	41 cm	84 cm	Not known	79 cm	100 cm
No. of source locations	Not known	Not known	Not known	685	575	937	625
Range in Elev.	$\leq 50^\circ$	$-80^\circ - 80^\circ$	Not known	$-90^\circ - 90^\circ$	$-82.5^\circ - 82.5^\circ$	$-40^\circ - 90^\circ$	$-45^\circ - 225^\circ$
Resolution in Elev.	Not known	6.15°	10°	5°	7.5°	5°	7.5°
Range in Azim.	$\leq 90^\circ$	$-130^\circ - 130^\circ$	$-90^\circ - 90^\circ$	$-90^\circ - 90^\circ$	$-90^\circ - 90^\circ$	$-180^\circ - 180^\circ$	$-90^\circ - 90^\circ$
Resolution in Azim.	Not known	2°	10°	5°	7.5°	10°	7.5°
Sampling rate	Not known	Not known	Not known	500 kHz	250 kHz	97.65625 Hz	97.65625 kHz
Input signal	Pure tone 25,45,65,85kHz	Pure tone 5−125 kHz	Pure tone 5−80 kHz	linear FM, 2ms	Random noise burst,1s	frequency chirp	11^{th} order MLS
SNR	Not known	Not known	65 dB	Not known	50 dB	Not known	Not known
Head positioning	Visually	Visually	Visually	Photogragh, Two lasers	Visually	Visually	Three lasers
Microphone	B&K 4135	B&K 4138	B&K 4138	Not known	B&K 4135	B&K 4182	B&K 4182
Microphone-ear connection	Plastic tube adaptor	Plastic cone	Directly	Metal probe	Metal tube	Silicon tube+ probe tube	Probe tube
Loudspeaker	Not known	Custom-built	Ribbon tweeter	Custom-built	Messring, Polaroid 8665	Sony, SS-TW 100ED	Morel, MDT-20
Frequency analysis range	25,45,65,85kHz	5k−125kHz	10k−80kHz	10k−90kHz	25k−95kHz	up to 50kHz	2k−40kHz

Table 4.2: HRTF measurement in current study

Specification	**Current study**
Subject	Bat
Head diameter	30-40 mm
Position of the receiver at the ear	Entrance to the ear canal
Distance	90 cm
No. of source locations	48
Range in Elev.	n/a
Resolution in Elev.	n/a
Range in Azim.	$-180^\circ - 180^\circ$
Resolution in Azim.	7.5°
Sampling rate	500 kHz
Input signal	256 k pink noise
SNR	50 dB
Head positioning	Two lasers
Microphone	B&K 4939
Microphone-ear connection	Metal tube or right angle connector (B&K EU-4000)
Loudspeaker	Ultra Sound Advice S56
Frequency analysis range	5−100 kHz

Table 4.3: Data processing method

	Interaural level difference	**Interaural time/phase difference**
Jen and Chen [1988]	Sound pressure level measurement for each frequency	n/a
Obrist et al. [1993]	Bandpass filtered	n/a
Fuzessery [1996]	Peak-to-peak gain analysis	n/a
Aytekin et al. [2004]	Rectangular window, 2048-point FFT	1/3-octave band 2^{nd} order Butterworth filter banks, envelope cross-correlation
Firzlaff and Schuller [2004]	Retangular window, 512-point FFT	n/a
Maki and Furukawa [2005]	Hamming window, 512-point FFT	Zero-padding, 32768-point FFT, minimum phase method
Koka et al. [2008]	Hanning window	Hilbert transform, envelope cross-correlation
Current study	Hanning window, 8192-point FFT	Hilbert transform, envelope cross-correlation

4.3 Current measurement system response

4.3.1 Measurement set-up

(1) Speaker and Microphone set-up

All measurements were conducted in the 'small anechoic chamber' in the Institute of Sound and Vibration Research (ISVR). The signal was generated from a speaker (Ultra Sound Advice S56 10 kHz-200 kHz) and amplified (S55A Amplifier 18 kHz- 300 kHz, ± 3 dB). The signal received at the microphone (B&K Type 4939) was amplified (B&K Type 2670 preamplifier 4 Hz-100 kHz and B&K 2690 conditioning amplifier with 140 kHz upper limit) and transmitted to the A/D (analog/digtal) converting data acquisition card with a sampling rate of 500 kHz. Measurements were controlled by a computer located in a room adjacent to the chamber. The two rooms are connected by a box-shaped tunnel for the connection of the cables.

(2) Near-field and far-field

A near-field can be defined as a acoustic radiation field which is close to an acoustic source such as a loudspeaker, therefore the frequency characteristics are affected by the radius of the loudspeaker. The measurement should avoid receiving the response in the near-field because near-field responses do not fall off in amplitude inversely proportional to the distance. This is due to the fact that radiation of the sound energy in the near-field is not uniform. To examine the near/far-field responses, a 256,000 pink noise signal was generated. Then the sound pressure level in arrangements of microphone only and of microphone with the custom-made tube, were measured. This was repeated for various distances from 0.05 m to 2 m. Figure 4.4 plots the relative sound pressure level against the distance when the measurements are made for microphone only and microphone with the tube. For both cases, a reduction appeared when the distance was greater than 10 cm. To help define near field and far field, the 6 dB decaying plot is also simulated for each case by setting the initial condition of sound pressure level at 10 cm as same as the measured value. This additional simulated plot helps for better observation of 6 dB-decaying tendency. As a result, the measured data is shown to have errors compared to the simulated result within ±1.7 dB for the case of microphone only, and ±1 dB for the case of microphone with the tube.

(3) Sine wave test

A sine wave pure tone was generated at different frequencies in 1 kHz steps from 5 kHz to 100 kHz to examine distortion in the sound reception. For each measurement, the signal duration was set to 0.1 seconds to allow more than a few cycles to be recorded

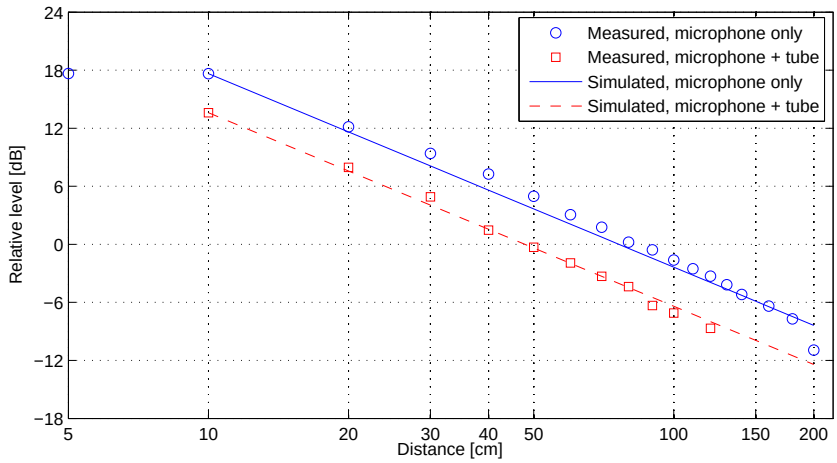


Figure 4.4: Examination of near-field and far-field: relative sound pressure level measured in various distances. Simulated plot helps to understand the 6-dB reduction (see text for details)

for each frequency. The signal was transmitted and recorded through the microphone which was located 90 cm from the speaker. Figure 4.5(a) shows the wave forms of the generated and recorded signals for 5 k, 10 k, 20 k, 40 k, 60 k, 80 k, 90 k and 100 kHz. The effects of noise and distortion in the time domain are shown particularly at 5 k and 100 kHz. This seems to be due to the relatively poor SNR at those frequencies. The whole frequency region was evaluated in terms of the magnitude response as shown in Figure 4.5(b). This is calculated using the difference between the relative amplitude of the transmitted and recorded signals.

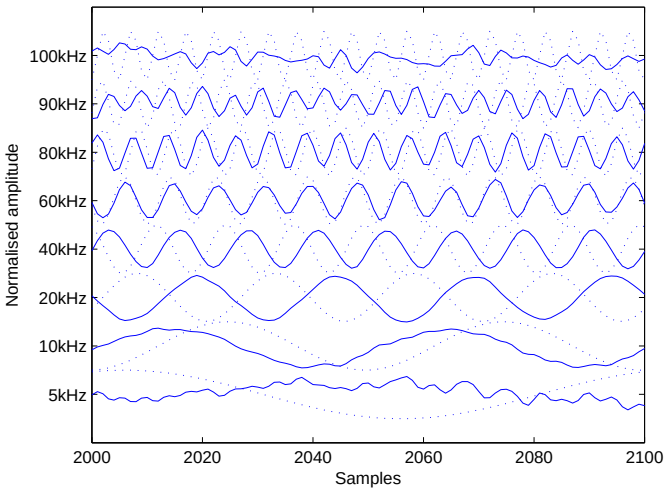
The phase response was also calculated and, to do so, 1000 samples were extracted from both the generated sine wave and the raw recorded data in the same time sequence. The cross-correlation function was then evaluated between them. For each frequency, the sample location where the cross-correlation function has a maximum value was detected within one cycle length of the corresponding frequency. Then the number of samples which indicate the phase delay, was transformed to time data ($\phi(\omega)$) and divided by the angular frequency (ω) to evaluate the radian value (τ_ϕ) for each frequency as in Eq 4.2

$$\tau_\phi = -\frac{\phi(\omega)}{\omega} \tag{4.2}$$

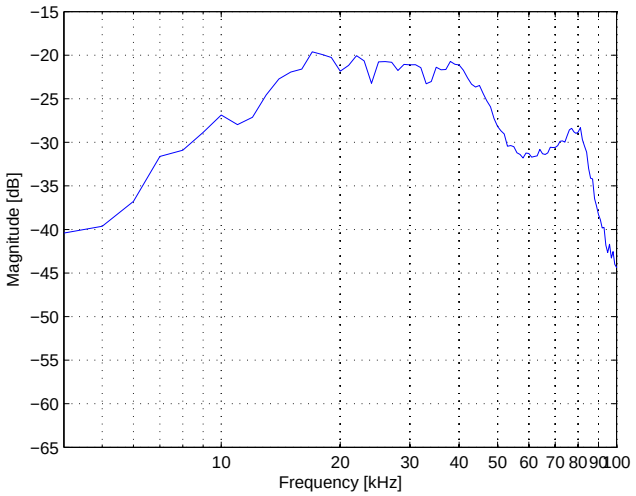
As the results shown in Figure 4.6(a), the phase delay decreases linearly in general but a phase jump occurs at a few frequencies. This phase jump was investigated and cor-

rected by subtracting 2π repeatedly from the values of those frequencies until all phase jumps were removed. It seems to be due to the fact that the recorded data does not begin with the point when the sound is sent out, but rather with the point when the sound is received at the microphone after it transmits 90 cm away from the speaker. As phase delay depends on the distance over which the sound wave travels, the delay which contains the multi-cycles of a given frequency might have happened. This is especially true in the higher frequencies which have shorter wavelength. Therefore, the phase jump with the multiple of 2π is induced and can then be corrected by subtracting 2π continuously until the unwrapped phase delay varies in same direction as the frequency increases. The corrected phase response is shown in Figure 4.6(b) and the theoretical value at a given distance is also plotted. For the theoretical value, the distance (r) of 88 cm and the speed of sound (c) of 344 m/s is employed to represent the current system and calculated using Eq 4.3. The result appears to have good agreement with the measured data. It is concluded that the current measurement system has a linear phase delay over the frequency range of 5 kHz to 100 kHz. The response shows different magnitudes depending on the frequencies. This response is considered to be consistent throughout the measurements of the HRTF and it can be removed by a free-field response calibration which is discussed in section 4.3.2. It is also shown that the SNR should be over approximately 20 dB for the signal in maximum amplitude, to enable the distortion-free analysis over the frequency range from about 5 kHz to 100 kHz.

$$\tau_{\phi,theoretical} = \frac{r\omega}{c} = \frac{r2\pi f}{c} \tag{4.3}$$

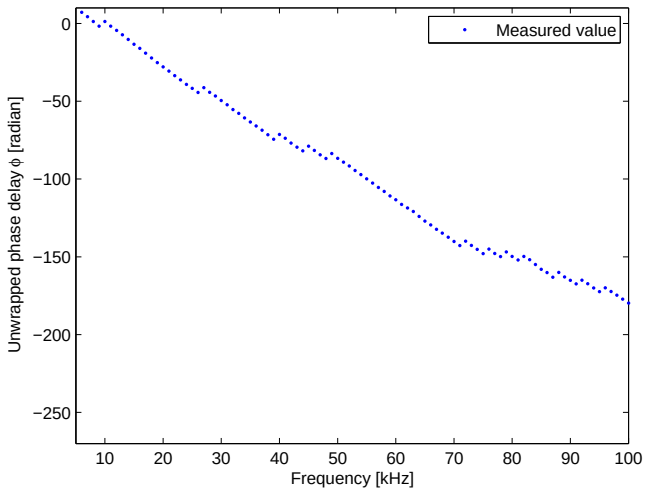


(a)

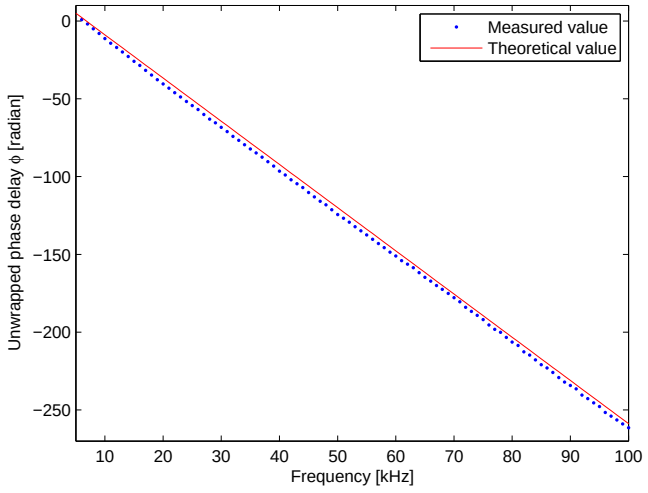


(b)

Figure 4.5: (a) Pure tone generation and recorded signal, dotted line: transmitted signal, solid line: received signal (b) magnitude response of the sine wave test in different frequencies from 5 kHz to 100 kHz



(a)



(b)

Figure 4.6: Phase response of the sine wave test (a) after cross-correlation (b) after correction of phase jump appeared in '(a)' (see text for details)

4.3.2 Free field response

The free-field response was measured for three different microphone arrangements. These allowed the characteristics of the custom-made tube to be examined. Figure 4.7 shows the three arrangements used: a) microphone only, b) microphone with tube angled (approximately 30° to the horizontal), c) microphone with tube straight.

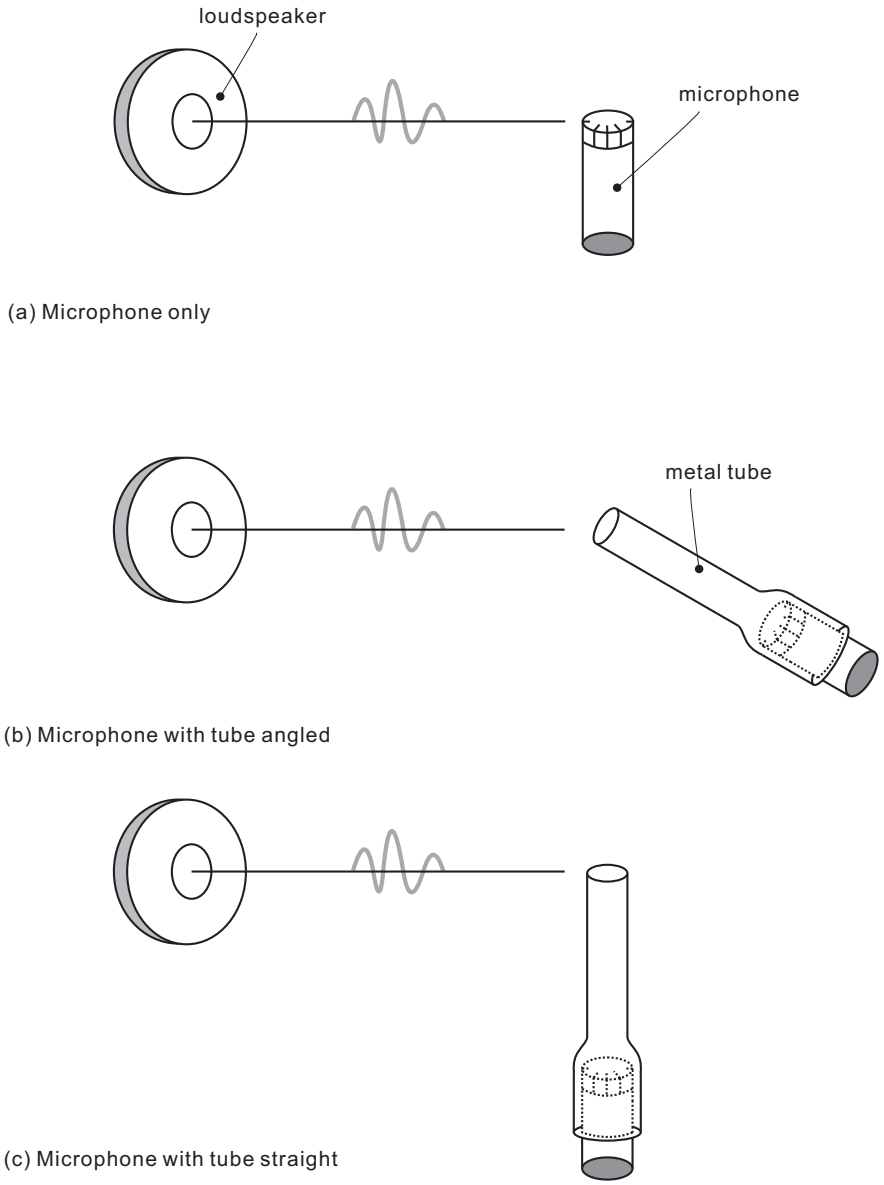


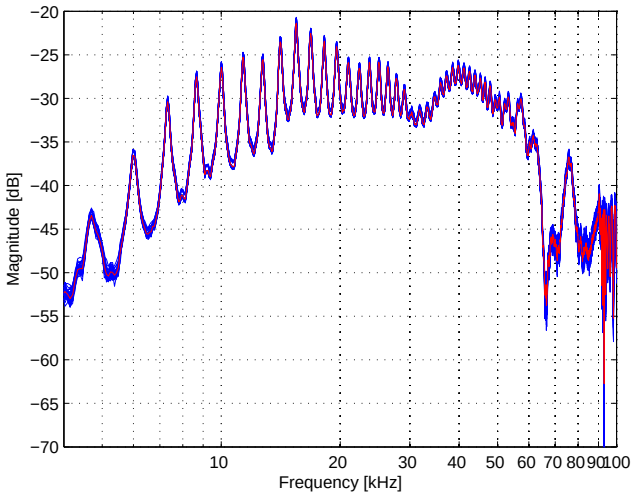
Figure 4.7: The positioning of receiver: (a) microphone only, (b) microphone with tube angled, (c) microphone with tube straight

A pink noise signal with 256,000 samples in length was generated from the loudspeaker with a 500 kHz sampling rate. The measurement was repeated 40 times for each ar-

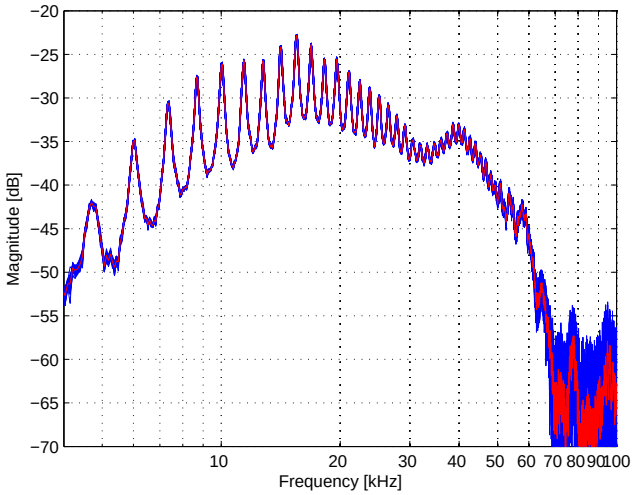
rangement to examine the repeatability in the reproduction and reception of the signal. A 1 second break was employed between the repetitions to avoid any overlapping or leaking of the received signal into the data acquisition card by the previous recording. In each recording, the pink noise was continuously generated 18 times to average the received signals, as 18 times averaging has shown to provide reasonably good quality of the signal. The first and last recordings were discarded to remove the effect of the data acquisition card's memory storage effect from the previous input. As a result, the data averaged over 16 cycles was finally represented. The impulse response was calculated automatically after the recording by an algorithm written in the Matlab environment. This algorithm performs the deconvolution of the received signal with the generated signal which is looped back to the data acquisition card. The post-processing is then applied to the impulse response by a rectangular window function of 200 samples in length around the peaks and zero-padded regions, from 201 to 8192 samples. A 8192-point fast Fourier Transform (FFT) was then applied to the post-processed impulse responses to determine the magnitude response in the frequency domain. The current analysis provides approximately 61 Hz of frequency resolution. From the magnitude response, the system reproducibility was evaluated by detecting the variation of the magnitude over 40 repetitions. The maximum and minimum amplitudes in each frequency bin were detected and the difference between them was used to describe the error. The distance between the microphone and speaker was set to 90 cm. The measurement height was set to 57 cm for the microphone only and 61 cm for the microphone with tube. The height of the centre point of the speaker was set to the same height as the microphone or the hole in the tube in each case. The floor of the anechoic chamber was covered with the foam wedges identical to those on the walls.

First, the raw response of the microphone with tube before the post-processing appeared to have multiple peaks in the frequency response as shown in Figure 4.8. A series of peaks are shown and the amplitude of those peaks is measured to be 10 dB. It appears to be induced by the resonance of the tube when the sound travels down the tube to the microphone. In addition, the difference between the straight tube and the angled tube has shown that the response from the angled tube has a higher magnitude in the frequency region above 30 kHz. The case with the straight tube has shown a steep falling-off above 40 kHz. Yet the arrangement of the angled tube shows a reasonably flat response within 5 dB variation up to 60 kHz, but has a prominent reduction in magnitude above 60 kHz. It is obvious that the degree in which the tube is angled dominates the result of the free-field response because it changes the angle of the sound incidence. The results after the post-processing are shown in Figure 4.9. The

microphone only arrangement shows the variation in the reproducibility less than 3 dB for the frequency region from 4 kHz to 100 kHz. It also agrees with the result from the sine wave test. For the two arrangements of the microphone coupled with the tube, the resonance effect has been shown to be removed after the post-processing. The case of the angled tube has shown the variation in reproducibility less than 3 dB for the whole frequency region studied. However, the frequency response shows a lower magnitude above 60 kHz compared to that of the microphone only set-up. The arrangement of the straight tube showed less magnitude than that of the angled tube in general at the frequency response. It also appeared to have large and fluctuating variation in reproducibility above about 70 kHz.



(a) Microphone with tube angled



(b) Microphone with tube straight

Figure 4.8: The frequency response of system with the metal tube before the post-processing. 5000 samples in each impulse response is extracted. The results show all responses from 40 repeated measurements and the averaged response. (a) microphone with tube angled (b) microphone with tube straight

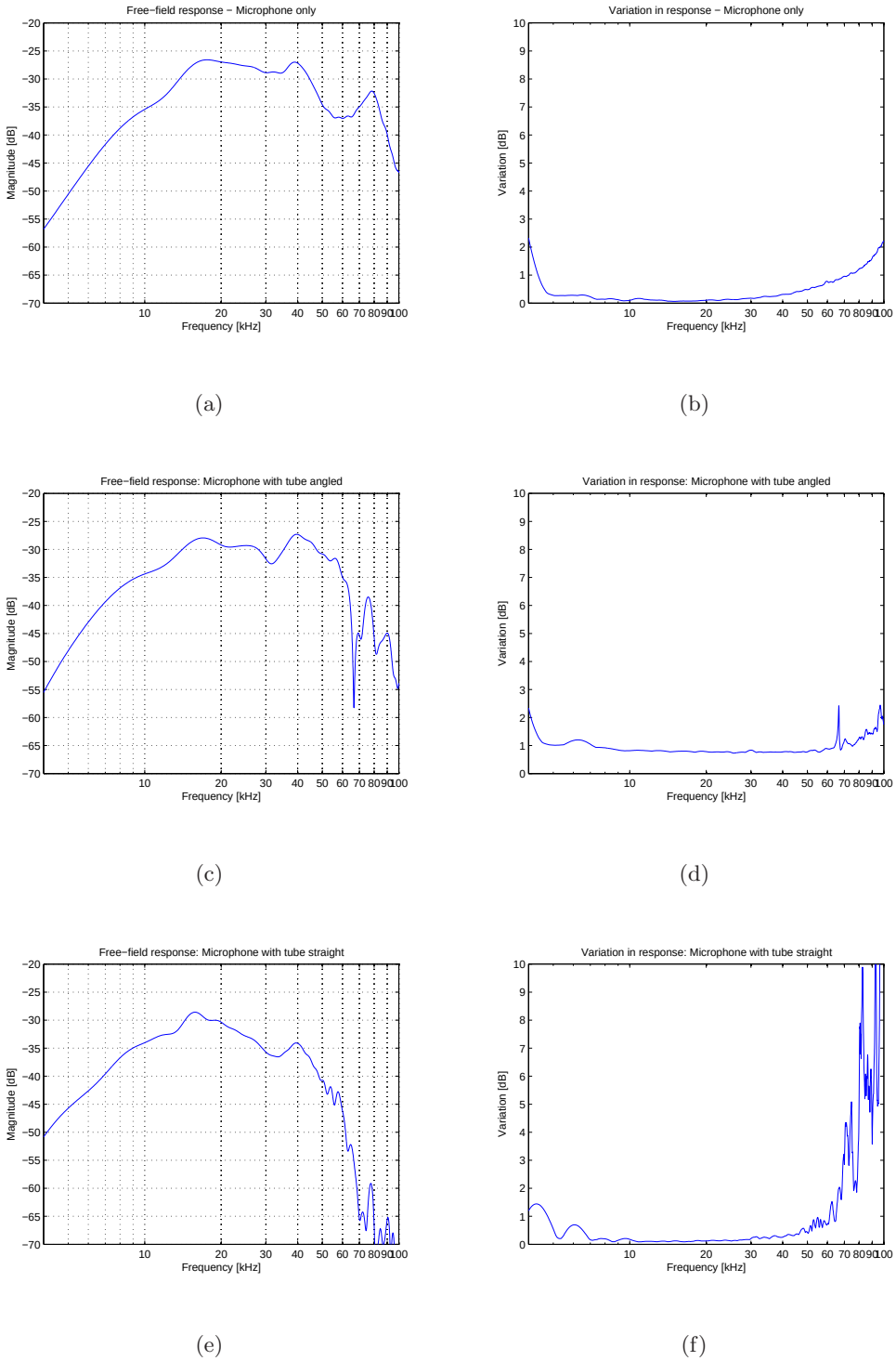


Figure 4.9: Free-field response after post-processing with 200-point Hanning windowing and zeroing for three cases of microphone positioining. (a)(c)(f) free-field response (b)(d)(f) variation between the repeated measurements (difference between maximum and minimum values)

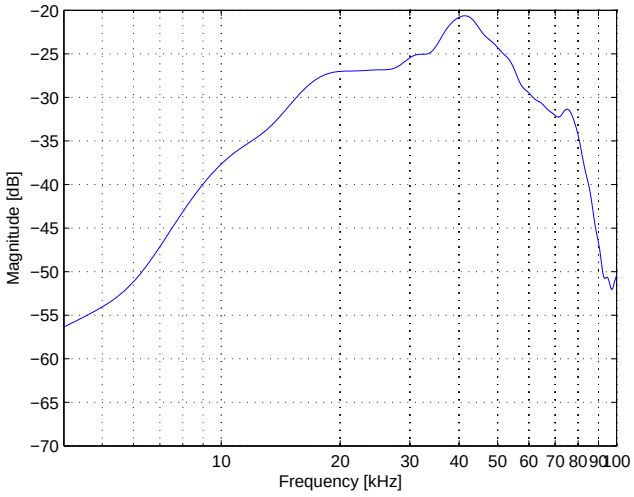
4.4 HRTF measurement with the bat-head cast

4.4.1 Microphone positioning within the ear

In the small-headed animal HRTF measurement, it is an important issue to select the correct microphone set-up as shown in section 4.3. In addition, the microphone should be inserted carefully in position in the head around the ear where the measurement is taken. The position of the microphone in the ear can be varied as four positions are possible: 'entrance to the ear canal', 'inner position in the ear canal', 'close to the ear drum (tympanic membrane)' and 'where the ear drum is located'. This measurement aims to fix the microphone at the position of 'entrance to the ear canal'. Therefore the HRTF result has been limited to include only the effect of the head and pinna rather than the effect of the ear canal. Because of the small dimensions of the head, the positioning of the microphone's shaft is critical as it is not easy to insert it perfectly inside of the head. Any incorrect positioning could disturb the measurement. Using a microphone extension tube as described in the previous section is one alternative method. The end of the metal extension can be inserted reasonably accurately at the position of the outer ear and the microphone can then be supported outside of the head. However, as described before, this approach is limited due to the fact that the tube response limits the acceptable frequency region of the measurement down to 70 kHz.

Another alternative method is to use a right-angled microphone. The right-angled connecter can be screwed to the microphone so that it changes the angle by 90 ° but does not change the frequency response of the original microphone. It enables the microphone shaft to be held straight during the measurement so that it reduces the positioning errors. The right-angled microphone and its frequency response are shown in Figure 4.10. The frequency range from 5 kHz to 100 kHz shows the magnitude within 35 dB and the high frequency response above 70 kHz appears to roll off quickly. The characteristics match reasonably well to the response of the microphone before the right-angle connecter is attached. The difference between them is due to the direction of microphone's membrane; the top of the microphone was aimed towards the speaker when the right-angled microphone was used. Therefore, the HRTF measurements have been carried out with the right-angled microphone. The experiment was repeated twice to investigate the repeatability of the measurement.

(a)



(b)

Figure 4.10: Right-angled microphone (a) photograph (b) frequency response

4.4.2 Material used in the experiment

A soft head-cast of the Egyptian fruit bat (*Rousettus aegyptiacus*) was used to measure the HRTF. The Egyptian fruit bat is the only species to use echolocation among Megachiroptera (Megabat) species. It produces a broadband echolocation sound by suddenly releasing the tongue upward away from the floor of the mouth. Each individual call lasts for a duration of 1 to 5 ms. The dimension between the two ears is approximately 3-4 cm. The bat-head cast used in this study is shown in Figure 4.11. The bat-head cast consists of a 'skin' which was taken from the cast of a real bat head, and it was provided by Dr Dean Waters of the Department of Biology at University of

Leeds through the BIAS (Biologically Inspired Acoustic Systems) consortium. Modelling clay was used to fill head cavity of the cast to maintain the shape of the head. Lacquering has been applied inside and outside the cast several times to keep the shape fixed.

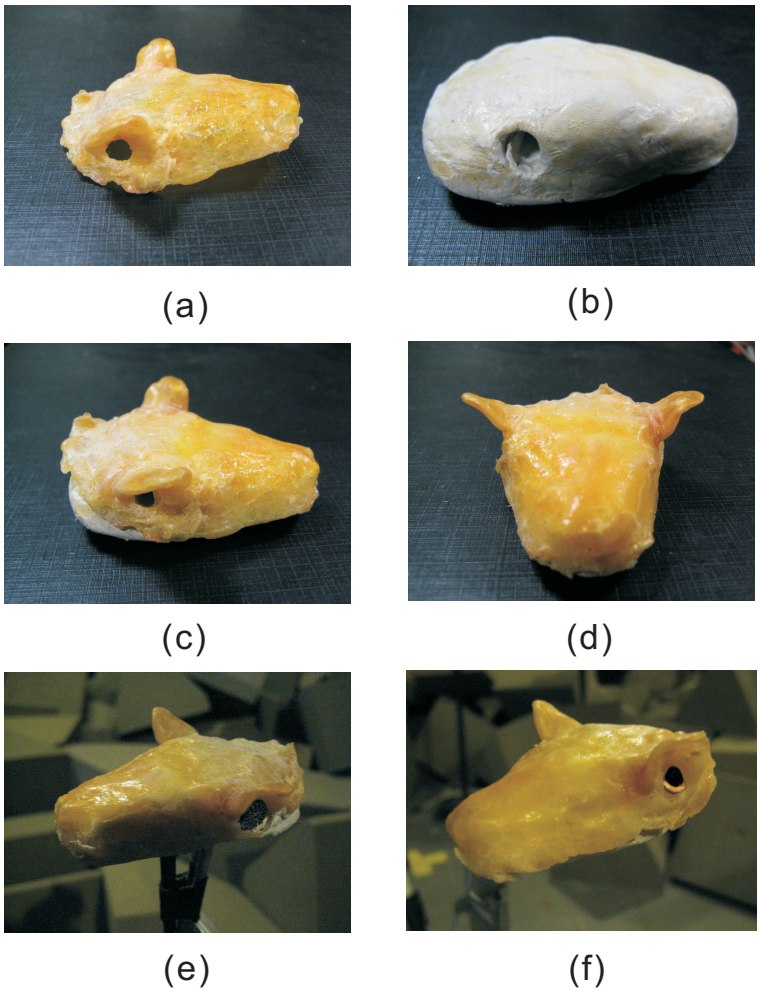


Figure 4.11: The soft bat-head cast. (a) latex skin (b) modelling clay (c) view from the side (d) view from the front (e) side view when the right-angled microphone is inserted (applied in current study) (f) side view when the custom metal tube is inserted (not used in current study)

4.4.3 Measurement and data processing

The HRTF measurements were taken in the small anechoic chamber (ISVR) and the measurement set-up was described in section 4.3. The measurement apparatus is shown schematically in Figure 4.12. The 10 cm long microphone shaft was fixed into the bat-head cast and then mounted on a metal stand. That was fixed on top of a turntable to record the microphone response at varying angles. The angle between the head and

speaker was set by rotating the turntable according to the azimuth scale fixed to the bottom of the turntable. The centre of the speaker was set at a distance of 90 cm away from the mid-point between the two ears of the bat-head cast. The height to the centre of the speaker was set to 60 cm above floor level, and the height to the microphone was also set to 60 cm above floor level to position the speaker and microphone in line. To position the centre of the bat-head cast and the centre of the speaker in line, a laser beam was projected over the experimental rig and the positions were adjusted before each measurement. An example of the projection of the laser beam is shown in Figure 4.13. This central line adjustment was made when the bat-head cast was located at 0° , 90° , 180° azimuth and was repeated until the three positions were corrected. The view of the measurement set-up in the small anechoic chamber is shown in Figure 4.14.

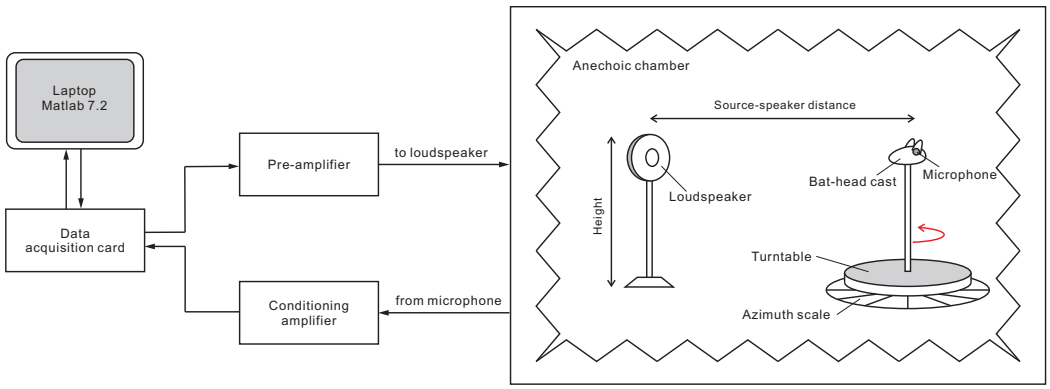
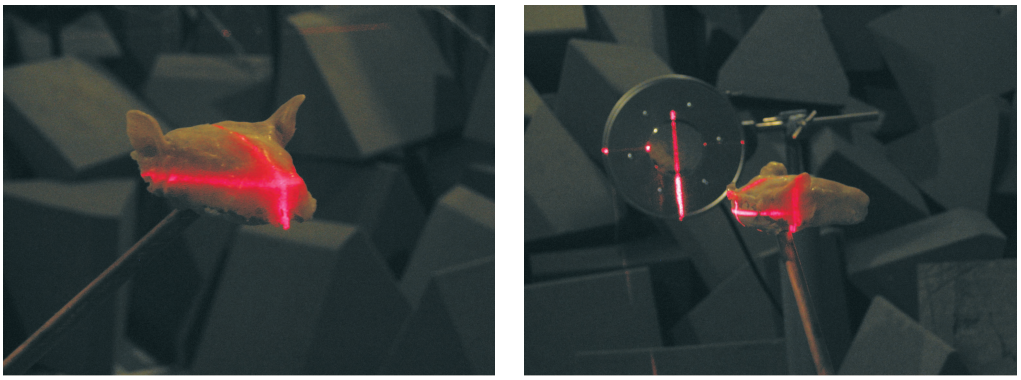


Figure 4.12: Measurement apparatus

The free-field measurement was made with the same microphone as used in the bat-head cast measurement but the response was recorded without the bat-head cast in place. The position of the microphone was re-adjusted using the laser beams to locate it in the same position as the middle of two ears of the bat-head cast. The measurement was made at a distance of 90 cm and height of 60 cm. As shown in the Figure 4.10(b), the frequency response is not flat but it can be easily removed from the measured transfer functions assuming its own shape is consistent during the measurement. The effective frequency range seems to be from 5 kHz to 95 kHz.

The HRTF measurement was made separately for each ear. A complete set of measurements were made at every 7.5° step in the azimuth, 48 positions in total, in the horizontal plane for the left ear and then the right ear. After the left ear was measured for all azimuths, the microphone was completely disassembled from the bat-head cast. Then, the microphone was fixed at the position of the right ear and the measurement



(a) (b)

Figure 4.13: Laser levels used to adjust the center position in the interaural axis (a) 0° adjustment. The laser beam crosses the mid-point between two ears. (b) 90° adjustment. The laser beam crosses the hole at ear ear.

Figure 4.14: Measurement set-up in the anechoic chamber (ISVR). The floor of the chamber is covered with the sound absorbing materials.

for all azimuths were retaken. The process of disassembling the rig was intended to examine the repeatability of the measurements. All measurements were repeated twice.

A pink noise signal with 256,000 samples was generated and the received signal was deconvolved from the original signal to produce the impulse response. The gain in the signal generating program, as well as the gain in the pre-amplifier of the speakers, were adjusted until the maximum SNR of the impulse response was achieved above 50 dB. Accordingly, the gain of the conditioning amplifier connected to the microphone for the

signal reception was adjusted so as not to be overloaded.

To analyse the magnitude response or the interaural level differences (ILDs), all impulse responses were processed with a 250-point rectangular window around the maximum peak response and the windowed signal was zero-padded in the tail with 7942-samples to produce the post-processed 8192 samples. Then 8192-point FFT was applied which corresponds to the frequency resolution of approximately 61 Hz in the analysis. The frequency analysis was limited to between 5 kHz and 100 kHz. The same processing was applied to the free-field response and the magnitude of the HRTF at each location was obtained by dividing the amplitude spectrum at that location by the free-field measurement response. On the other hand, the interaural time differences (ITDs) were calculated from the cross-correlation of the envelopes obtained from two ears measurements using the Hilbert transform in the time domain. The free-field equalisation was not applied in the processing of the ITDs to exclude the inevitable errors being induced by the centre positioning procedure as the sensitivity of the positioning error can be noticeable in the ITDs.

4.4.4 Results

(1) HRTFs analysis and contour plots

The magnitude responses of the two sets of the measurements are shown in Figure 4.15. In human studies, it has been known that the equalised HRTF at 0° converge at approximately 0 dB as frequency decreases, since the presence of the head hardly affects the sound field in the low frequency range. The 'low frequency' is conventionally less than 1 kHz in human studies. In this study, the same effect was shown in the measured response (b) from the sound source at 0° azimuth in the frequency range less than 10 kHz. In addition, the same effect was shown in the different azimuths of 90° , 180° and -90°, hence the sound in this frequency region seems to reach both ears regardless of the presence of the bat head cast. It is noted that the frequency response below 5 kHz are not reliable as the convergence to 0 dB is not fully achieved in this region. It is considered to be due to the limited frequency range in the free-field response, as discussed earlier. Another expected feature in the HRTF is the pinna notch, and this is shown in the results at around the 80 kHz region. The pinna notch is not very clear in the 180° azimuth where the sound source is coming from behind the head. It seems to be due to the fact that there is no direct path to the receivers at this azimuth.

From the two repeated measurements, the results can be compared to analyse the potential errors in the HRTF of the bat-head cast. The difference between the two

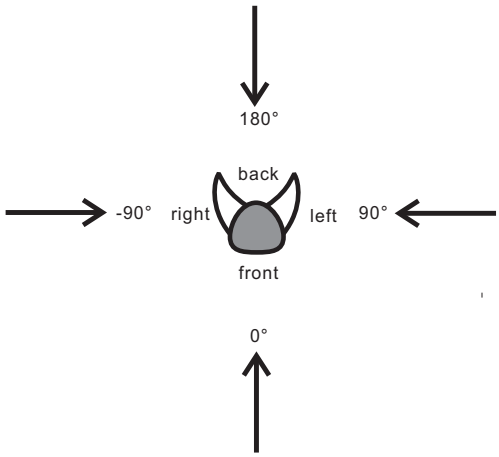
measurements for each ear is shown to be generally less than 3 dB over the frequency range from 5 kHz to 20 kHz. Larger differences are shown in the higher frequencies but it was generally limited to the cases which there are no evident direct path between the sound source and the receiver. This is clear in 90° and −90° azimuths as the differences between the results from the two measurements are consistently low in the higher frequencies for the receivers in the direct paths (the left ears at 90° and the right ears at −90°). However, the opposite ears, which are located on the opposite side to the direct paths, showed larger differences between the two sets of measurements in the high frequencies. This implies that the high frequency response is vulnerable to the positioning only when the indirect path is dominant between the sound source and the receiver. Accordingly, the peaks and notches shown in the response from the indirect paths do not match clearly between the two measurements. Therefore, careful consideration for the analysis of the spectral peaks and notches is recommended in the HRTF analysis of this study. It is also noted that the left and right sides of the bat-head cast are not symmetrical as the frequency responses do not exactly match.

The spectral contour plots of the HRTFs measured in all azimuths provide several characteristics which explain the acoustical phenomenon around the bat-head cast. Four contour plots from two measurements at each ear are shown in Figure 4.16. The 8192-point FFT was applied to the data and the results presented in dB at a frequency range from 5 kHz to 100 kHz. The obtained data from the 7.5° azimuth steps is interpolated to represent a smooth contour. Each contour appears to have a different range of maximum and minimum points, and there is asymmetry and differences in the two measurements. Therefore, all contours are plotted over the range of 40 dB with different maximum and minimum absolute values to aid understanding of the characteristics in the same range of colours. Thus, the colour bar shown in Figure 4.16 represents the relative levels.

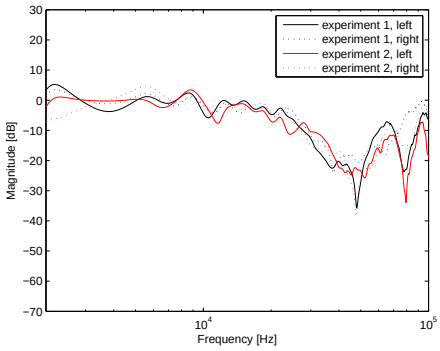
The general characteristics of the measured HRTF in all azimuths are shown to be consistent between the two sets of the measurements. They are observed to have similar patterns describing the principles of acoustics around the head. There are four major features shown in the contour plots. These are referenced as A, B, C and D in each figure at the same locations. The magnitude of the HRTFs relatively increases at high frequencies as the angle of the source increases towards the location of 90° in azimuth (A). It is obvious that more high frequency energy is received in this direction. This is due to the fact that the sound waves at high frequencies reflect back in the direction of the sound source when the sound source is located directly in front of the receiver. In

other words, the original sound waves and the reflected waves are combined in phase and they produce the high-frequency doubling effect.

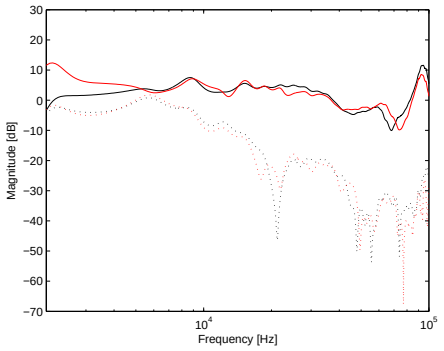
A large area of spectral notches was found for the sources located on the back of the bat-head cast (B). There is very low energy at high frequencies in this region as the travelling sound waves at high frequencies are shadowed by the head. Furthermore, the notch is somewhat extended and makes interference with the frontal region at around 70 kHz - 80 kHz (D). On the other hand, relatively higher energy of the low frequency is contained in the response for the source located in the back. This effect is most obvious when the ear is located directly opposite the source. Assuming the head is symmetrical, the theory shows that low frequency sound components travel around the head by diffraction and then combine in phase to produce the acoustical bright spot (C). As the bat-head cast used is not perfectly symmetrical the bright spot appears at less than $-90\,^{\circ}$ and the peak locates at a frequency of 10 kHz.



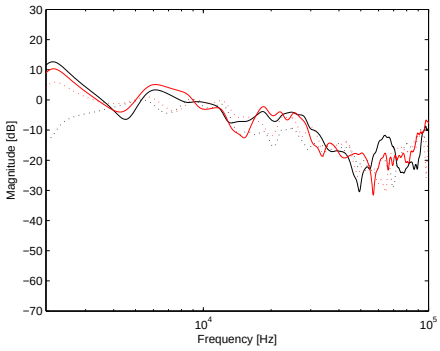
(a) source positions



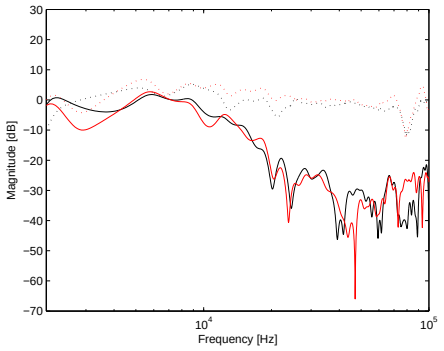
(b) 0 °



(c) 90 °

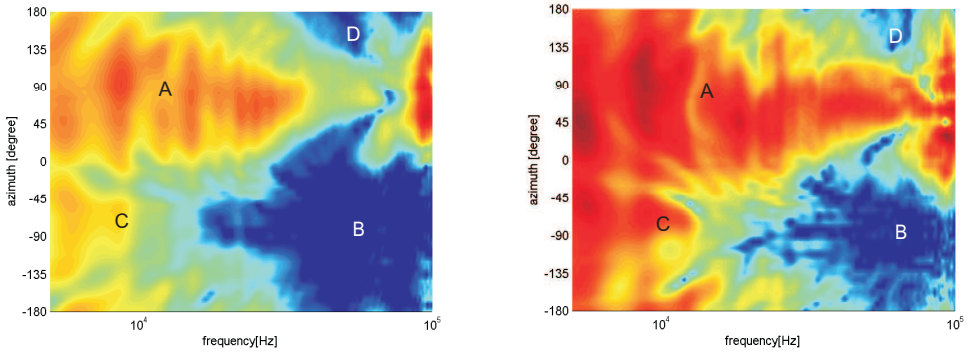


(d) 180 °



(e) 270 °

Figure 4.15: HRTFs in each ear for the source positions (a) at (b) 0 ° (c) 90 ° (d) 180 ° (e) 270 ° after post-processing. The results are shown for two experiments at each ear.

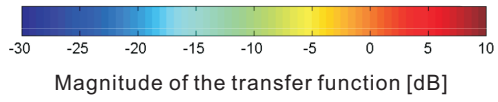


(a) Experiment 1, left ear

(b) Experiment 1, right ear

(c) Experiment 2, left ear

(d) Experiment 2, right ear



(e)

Figure 4.16: Contour plots of the HRTFs in each ear obtained from two sets of measurements. The current figure is interpolated using the data from 7.5° step. The 8192-point Fast Fourier Transform is applied and the frequency range is shown from 5000 Hz to 100 kHz. The magnitude of the transfer function at each point is represented by colour, as shown by the legend across the bottom of the figure. Four reference points (labeled A-D) are present on the contour plots (see text for details). A: High frequency doubling effect, B: Head shadowing effect, C: Acoustic bright spot, D: Extension of interference pattern to the ipsilateral hemisphere.

(2) ILDs

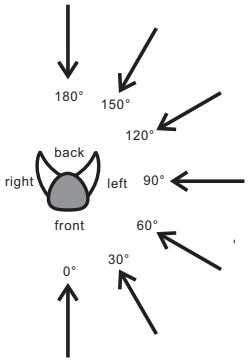
The ILDs are calculated for each measurement set. A few results from different azimuths are plotted in Figure 4.17 and Figure 4.18. The ILDs are identified near 0 dB in the median plan and the ILDs of the low frequencies from 5 kHz to 10 kHz tend to stay near 0 dB or increase up to about 10 dB at maximum as the source moves towards the lateral positions. The ILDs in the high frequencies above 10 kHz vary dramatically compared to the low frequencies. Several peaks and notches appear as the source moves, and they are shown to be stronger at 60° and 90° where the direct paths of the sound travelling in the high frequencies are given from the source to the receiver. The maximum ILDs reach around 50 dB in the region of high frequency.

The differences between the two measurements in the ILDs analysis showed that the results are consistent at the low frequencies from 5 kHz to 10 kHz with the variation of less than 3 dB. However, the spectral peaks and notches above 20 kHz are irregular between the measurements, and the difference between the peaks in particular frequency is shown to be over 20 dB.

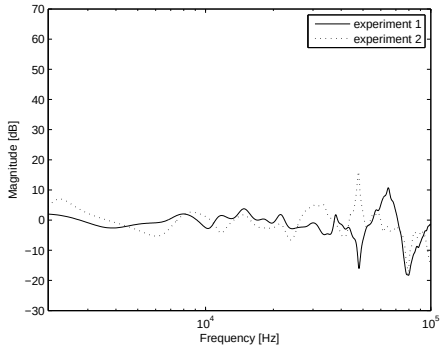
(3) ITDs

The ITDs are calculated for the two sets of measurements and plotted in Figure 4.19. The ITDs obtained were interpolated to represent the data in higher resolution in the azimuth between each step. The simulation of the ITDs (τ) based on the rigid sphere model was conducted for the head diameter (a) of 4 cm. The mathematical calculation for the simulation is described in Eq 4.4 (θ denotes the azimuth and c is speed of sound. $c = 344m/s$ is applied) [Middlebrooks and Green, 1990]. The range of the ITDs are shown to be from -150 μs to +150 μs for both simulated and measured data. The measured data seems to fit the sphere model in general. The ITDs increase as the source moves towards the lateral positions and decrease as the source moves towards the median plane. The negative values of the ITDs indicate that the right ear is closer to the source than the left ear. It is well known in the previous studies that the maximum ITD occurs at the position where the sound source is coming directly towards the ear (90° of azimuth). However, the maximum value of the ITD is not exactly shown in the source direction at 90° and it is shifted towards the smaller azimuth angle. The explanation is that the effect from the pinna is not included in the sphere model. Furthermore, the peak was shown to be less obvious for the second measurement and this seems to be due to the high sensitivity in positioning.

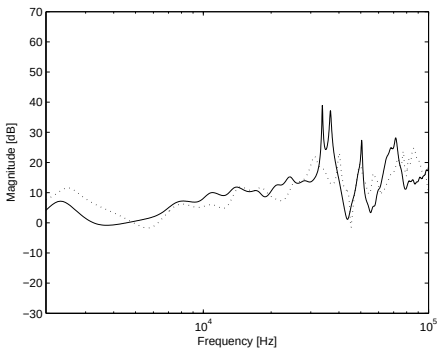
$$\tau = \frac{a}{c}(\theta + \sin\theta) \tag{4.4}$$



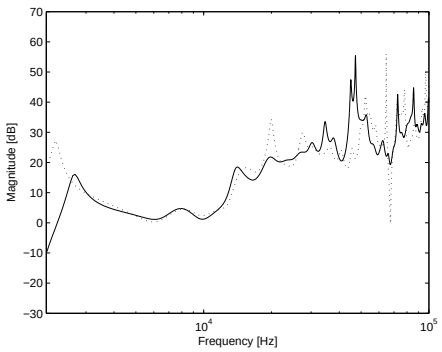
(a) source positions



(b) 0 °

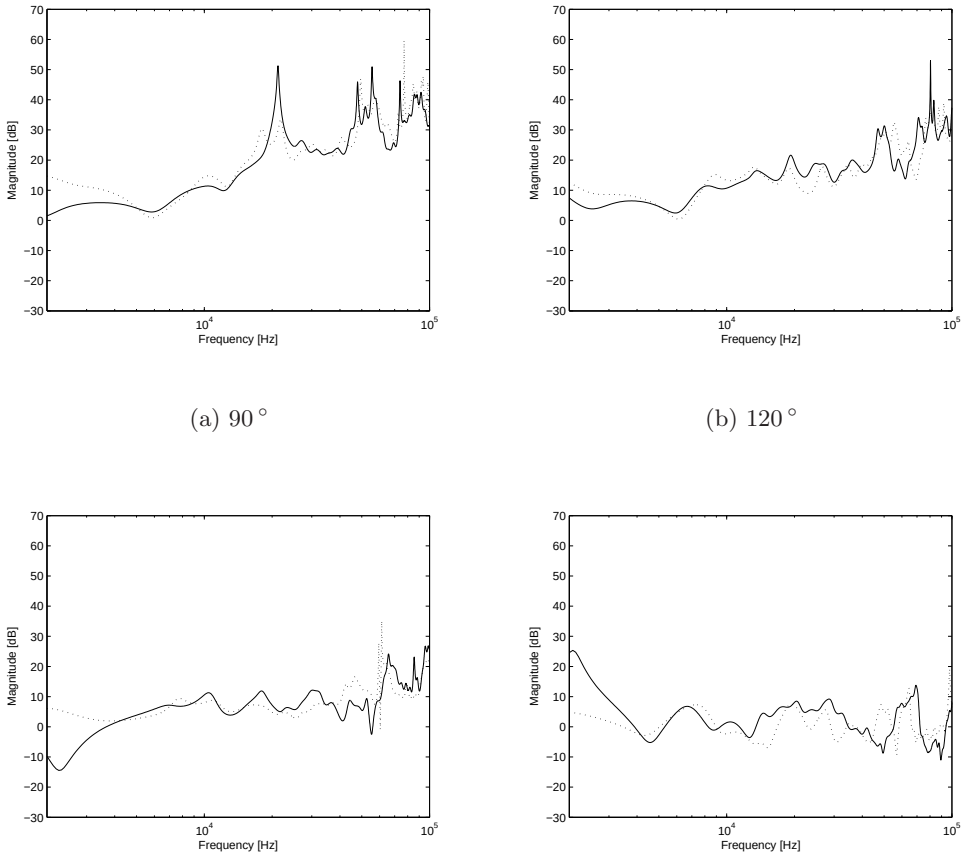


(c) 30 °



(d) 60 °

Figure 4.17: ILDs calculated from the two sets of measurements - 0 °, 30 ° and 60 °



(a) 90 °

(b) 120 °

(c) 150 °

(d) 180 °

Figure 4.18: ILDs calculated from the two sets of measurements - 90 °, 120 °, 150 ° and 180 ° (The source positions are same as described in Fig 4.17(a).

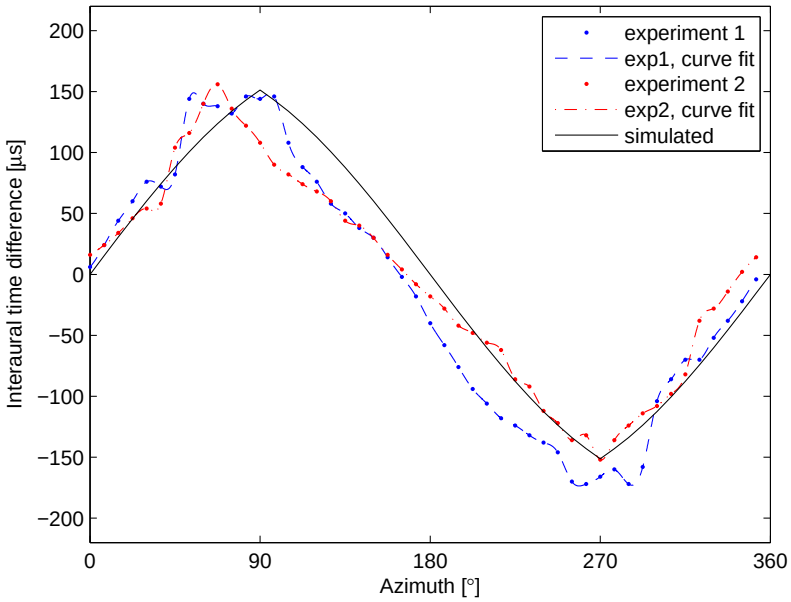


Figure 4.19: ITDs calculated from the rigid sphere model and the two sets of measurements

In general, the results from the two measurements showed less variation for the sound sources from the front and they matched reasonably well to the prediction of the sphere model. However, it has shown the variation up to around 50 μs for the sound sources from the back. Thus, the measured data is shown to be less reliable when the direct path is shadowed for the receivers in both ears.

4.5 Discussion

Based on the principle of the head related transfer functions and the review of previous studies regarding the HRTFs in small-headed animals, the current system to measure the bat-head cast was developed and evaluated to test reliability and repeatability.

This study suggests that one of the common issues is the insertion of the microphone into the small head and the subsequent deteriorated frequency response caused by using the extension tube attached to the microphone. In this study, a custom-made metal tube has been examined in two different positions; straight and angled. As expected, they showed resonant peaks in the frequency response. It has been possible to remove the peaks by windowing the signal in the time domain and discarding the reflected part which produces the resonance. Yet this post-processing reduces the magnitude of the frequency responses in the frequency region above 60 kHz. Furthermore, for the case of

a straight microphone, the reproducibility of the measurement was shown to be poorer at the frequencies above 60 kHz with irregular variation of 3 dB over the 40 repeated measurements. It implies that the angle of the tube affects the reproducibility as the response changes depending on the angle of incidence from the sound source towards the tube. It also means that careful consideration is required to produce exactly the same set-up in terms of the angle of tube towards the sound source, between the free-field measurement and the HRTF measurement. In the author's opinion, it is considered difficult to maintain a set-up when a tube is used, particularly in a study to measure different bat casts at the same time. Furthermore, as the position of the end of the receiver is also important and it should be fixed in the correct position in the ear which can be varied from 'entrance to the ear canal' to 'ear drum'. Therefore, it seems to be inevitable that the head and the angle of the receiver need careful adjustment. This study has shown that the measurement errors should be taken into account for high frequencies in particular when the tube is applied. It is noted that the high frequency response is important as it limits the system in terms of using the bat sonar frequencies.

Therefore, it is suggested in this study that the right-angled microphone is used which only changes the physical direction but not the frequency response. This study has successfully implemented this specific type of microphone in the bat-head cast and the HRTF measurements have been made reliably in the frequency region from 5 kHz to 100 kHz. The measurements were made twice at each ear separately. To investigate the repeatability, the microphone was completely disassembled from the bat-head cast and fixed again for each measurement and the centre position was readjusted.

The measured data was examined in the frequency domain and showed that the HRTF provided the change in the magnitude spectrum depending on the different azimuths measured. Also, the variation in frequencies above 10 kHz was larger than in frequencies below 10 kHz. It is interesting to note that common features have been found in the contour plots of the measured HRTFs and human HRTFs from previous studies [Brungart and Rabinowitz, 1999]. The high-frequency doubling effect, head shadowing, acoustic bright spot and the interferences shown in the high frequency regions are shown in the results and they are regarded as general characteristics of the measured HRTF. The same effect has been shown in the human HRTF and it seems likely that the scale has been shifted up to the frequency region of 10 kHz to 100 kHz in this study. In addition, the contour plots showed reasonably good repeatability in the measurements for the same ear, but it was obvious to see the asymmetry of the head by comparing the results from the two ears.

The ILDs and ITDs were also investigated. For ILDs, the results from two sets of the same measurements have shown reasonably good agreement in the low frequency region but the variation has appeared to be larger in the high frequency region. Some peaks appeared in the high frequencies, have not matched together and have not shown the same level of magnitude. The ITDs showed similar results compared to the simulated sphere model. They showed better repeatability and agreement to the model for the sources located in the front rather than at the back. In summary, it can be concluded that the ILDs and ITDs are more sensitive when the direct path between the sound source and receiver is head shadowed.

This study has shown the characteristics of binaural receivers fixed on the artificial bat-head cast. It is generally agreed that the binaural receivers provide the localisation information for different azimuths and this is similar to the results from human studies and other HRTF measurement in bats. In fact, the binaural characteristics can be divided into two: the effect from the head, and the effect from the pinna. Most studies of HRTF in bats have focused on the characteristics of the pinna and their contributions to the localisation using a specific type of echolocating signal. In contrast, the studies regarding the relationship between the presence of the head and its contribution to the echolocation task are relatively limited. The presence of the head can be actively used in the echolocation process by moving it up-and-down or side-to-side. Also the head approaching the target can be seen as an active motion shown in echolocation. The contribution of the head to echolocation is even supported by the psychophysical study in human echolocation. Rosenblum et al. [2000] has shown that human listeners can decide the location of an object using echolocation under experimental conditions. It is interesting to see that their performance improves when they were allowed to move around the object. Considering the similarity of the general characteristics in the measured HRTFs compared to humans, the motion of the head in the echolocation task is assumed to be similar for humans and bats. Thus, the role of the head and the binaural receivers on the echolocation motivates further study.

Chapter 5

Range-dependent Characteristics in the HRTFs of a Bat-head Cast

5.1 Introduction

In this chapter further development of the binaural receiver is described using a different form of bat-head cast from the previous chapter. Instead of using a soft bat-head cast, a hard and rigid cast was used and the HRTF was measured at various distances from the distal region to the proximal region. First of all, measuring HRTF at different distances raises the technical issue of setting up the standard distance for the measurement of small-headed animals. This issue was discussed in the previous chapter. More importantly, the measured data at different distances are expected to provide better understanding of binaural processing in echolocation and a more realistic understanding of the function of a bat's head during its flight while it is approaching a target. Sensing the target prey at distance and catching it at close proximity obviously means that the distance between the sound source (returning echo) and receivers (ears) varies over time. Therefore, it is expected that the binaural signals reaching the two ears change during echolocation and those signals will be affected by the response of the head, pinna, body and wings. As described in Chapter 4, this study only considers effects caused by the head and pinna, i.e the HRTF.

However, the analysis of binaural signals during the echolocation process is challenging due to the bat's nature of their active movements and adaptive change during the process. Echolocation is considered as an active sensing system that accounts for variations in the bat's behaviour during flight. There are more issues to consider such as flight pattern, signal characteristics, head and pinna movement. Previous studies, especially in biology, have shown what happens during echolocation by bats. It is well known

that bats experience search, approach and terminal stages during the echolocation process [Hill and Smith, 1984]. A typical sequence from search to terminal stage does not take longer than 1 second [Ulanovsky and Moss, 2008]. Therefore, it is expected to be difficult to analyse both sensing and movement in echolocation at the same time. Also, although the adaptive change in the echolocation has been observed by many researchers, the functional implication has not yet been clearly explained.

The current literature has shown the following details about the adaptive changes in acoustic and physical features observed during the bat's echolocation.

(1) Variation in echolocating calls

Griffin et al. [1960] has already observed that bats increase bandwidth and decrease duration during the approaching stage. Simmons and Stein [1980] have insisted that the bandwidth and average period of echolocation signals can determine the quality of target perception. It has been also reported that bats change the calls not only when they are attacking prey but also when they are approaching obstacles such as tree branches and ground [Ulanovsky and Moss, 2008]. Classic sonar theory explains that increasing bandwidth of the call improves the acoustical image of given objects while decreasing the call duration helps bats to avoid overlapping between the emitted call and received echoes. In addition, the other explanation about controlling bandwidth of the call is that adaptive change in the call design might compensate the systematic error induced during the flight. Increasing bandwidth at high frequency with is found in search calls. Siemers and Schnitzler [2004] have recently demonstrated strong correlation between the ability of prey detection and the respective bandwidth in echolocating calls in the 'search' stage. In particular, broadening of the bandwidth was obtained mainly by increasing the starting frequency. Again these results emphasise the role of wide bandwidth in better characterisation of the prey.

(2) Head and pinna movement

Head movement is also observed during the bat's echolocation but a recent finding has shown that the bat locks its head towards the target when the bat switches from search to approach mode (i.e., changes the rate of its vocal production) [Ghose and Moss, 2006]. It has been shown in several studies that some species of bats move their pinnae in correspondence with echolocating signals [Griffin et al., 1962]. The mobile pinnae are known to control the gain of the echoes. Holland and Waters [2005] have recently measured the HRTF of *Rousettus aegyptiacus* to demonstrate higher sensitivity in gain when its pinnae are directed forward rather than backward.

(3) Sound pressure level issue
Kick and Simmons [1984] suggested that the output level of the call is constant, but the sensitivity in the ear improves with the echo delay due to the fact that the echo level decreases as the distance increases. The intensity compensation or automatic gain control during the emission of calls was also studied by Hartley, and the results imply that bats might compensate the acoustical attenuation during the air transmission of echo signals [Hartley, 1992]. Boonman and Jones [2002] have shown that this behaviour is not dependent on the target, but it is global and the pattern is stereotyped regardless of the type of target. Surlykke and Kalko [2008] have argued that bats compensate intensity attenuation of different frequencies in their echolocation call as their findings showed that bats emit the higher intensities for the higher frequencies, which severely attenuate in the air. This evidence supports the idea that the echolocating bats might receive stabilised amplitude of the echoes which are independent of the frequencies.

Although the understanding of the relationship between the active echolocating motion and sensing have been made to some extent as described above, there are not many studies about the effect of binaural processing during flight. In other words, the functional role of binaural processing in echolocation has received only little attention except the general sound localisation at a fixed position. However, as there are major adaptive changes in the echolocating behaviour, it is hard to reveal the functional implications of binaural hearing in relation to every variation. Therefore, investigation of the adaptive change of binaural hearing under a specified and simplified condition is required. To do so, here the adaptive change in the echolocating calls and range dependence is taken into account to investigate the binaural processing in the echolocation. In other words, the binaural processing in this chapter will focus on the relationship between the frequency and distance which are modified by binaural information. The results are expected to provide a physical evidence about why bats change the frequency characteristics as they move towards a target. Of course, there are some explanations of adaptive change in calls as reviewed in Chapter 2, but there are still questions which have not been explained. For example, it has also been observed that some sonograms of echolocating signals show the extension of frequency bandwidth towards lower frequency in 'approaching' and 'terminal' stages [Hill and Smith, 1984; Simmons et al., 1995; Boonman and Jones, 2002], but the reason why the bats make this change has not yet been understood.

The study in this chapter simplifies the motion in flight to a change in distance and aims to investigate the relationship between the effect of distance on the frequency

characteristics when an artificial bat-head and pinna are presented. By using a bat-head cast, this change can be recreated by simply moving the position of the sound source relative to the hard bat-head cast. This chapter presents the monaural/binaural signal analysis at various distances of distal/proximal region to understand the role of different frequency band in echolocation.

5.2 Method

5.2.1 Hard bat-head cast

The 'hard bat-head cast' is used for the measurement in this chapter. Based on the basic of acoustics, smooth hard surfaces reflect sound waves whereas soft materials help absorb them. Therefore, the hard bat-head cast is expected to provide results with less reduction in sound energy. The soft bat-head cast was placed into a box of wet plaster and allowed to set. The soft bat-head cast was removed to reveal a void into which resin was poured to produce the hard cast. The hard bat-head cast and its dimensions are shown in the Figure 5.1. The size of the hard bat-head cast is the same as the soft cast, but hard enough to prevent shape deformation, which has been a problem with the soft cast. As shown in the top view, it has 4 cm of interaural distance. The length of the front is approximately 2 cm and the distance between front and back is approximately 5.5 cm. The pinnae of both ears are about 1.5 cm long but the angle of the two pinnae are not exactly symmetrical due to natural variation. Inside the cast, a hole is made at the entrance to each ear so that the right-angled microphone can be positioned inside. A large hole was also made at the bottom of the cast to enable the handling of the microphone insertion, then the upper part of the microphone shaft was encapsulated with rubber to be fixed inside the hole.

5.2.2 Apparatus and experimental set-up

All measurements were conducted in the 'small anechoic chamber' in the Institute of Sound and Vibration Research (ISVR). The walls, ceiling and the floor in the anechoic chamber are covered with sound absorbing wedge forms to prevent the reflections during the measurement. The output signal was generated from a laptop, amplified (S55A Amplifier 18 kHz- 300 kHz, $\pm$ 3 dB), and then sent to the speaker (Ultra Sound Advice S56 10 kHz-200 kHz). The signal received at the microphone (B&K Type 4939) was amplified (B&K Type 2670 preamplifier 4 Hz-100 kHz and B&K 2690 conditioning amplifier with 140 kHz upper limit) and transmitted to the A/D (analogue /digital) converter with a sampling rate of 500 kHz. Measurements were controlled by the

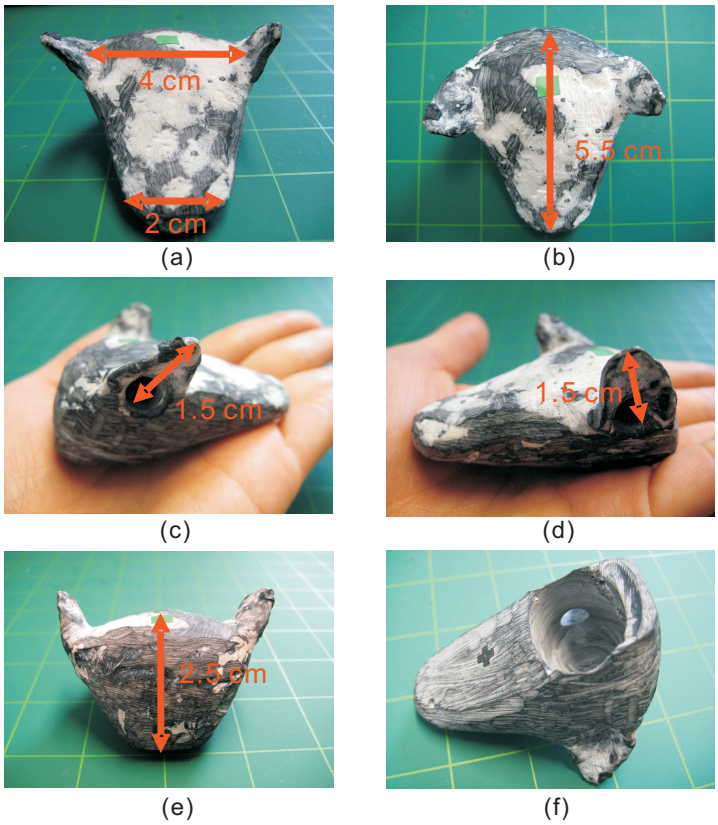
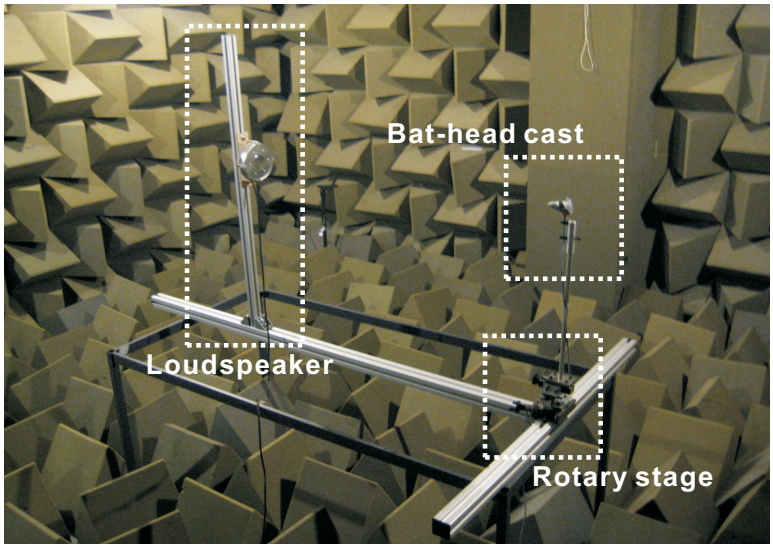


Figure 5.1: The new hard bat-head cast (a) front view (b) top view (c) view from right side (d) view from left side (e) back view (f) view from the bottom (showing where the microphone is inserted)

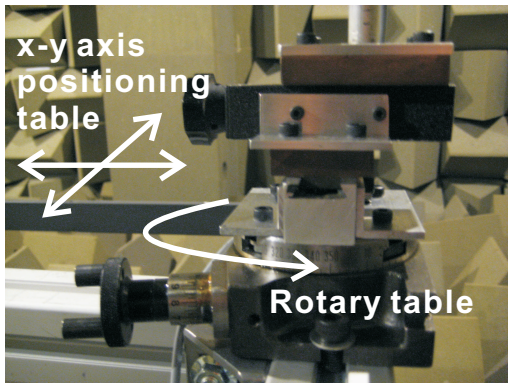
laptop located in a room adjacent to the chamber. These two rooms are connected by a box-shaped tunnel for the connection of the cables. Matlab 7.2 installed in the laptop was used to control generation, recording and processing of the signals. The measurement rig is shown in Figure 5.2(a). The microphone was inserted inside of the bat-head cast and the top of the microphone is fixed in pre-made hole which is positioned at the entrance of each ear. As the angle connector of the microphone rotates its angle by 90°, the top of the microphone can be positioned to face the hole at the ear horizontally. Figure 5.2(b) shows the bat-head cast with microphone fixed at the position of the ear. The microphone shaft was then fixed on the metal rod by using two clips. The metal rod was fixed on to the centre of the rotary stage which was designed specifically for the experiment. The rotary stage comprises the rotary table, which is adjustable in increments of 1° for a complete 360° of rotation, and the x-y axis positioning table, which enables fine adjustment for left /right and front /back positions. The set-up for the rotary stage is shown in Figure 5.2(c). The complete set of microphone, bat-head cast and the rotary stage was fixed at the measurement rig. The measurement rig was placed on top of a metal frame which is mounted on the floor

of the chamber. The height between the floor and bottom level of the rotary stage was approximately 70 cm and the bat-head cast was set 60 cm above the baseline of the rig. The reference position of the bat-head cast was the middle point between the two ears at the level of the microphone holes and it was defined as the centre of interaural axis for the current measurement. The speaker was fixed on the vertical metal beam and its centre was also positioned 60 cm above the baseline of the rig. The speaker was attached to the centre beam, which extends out from the bottom of rotary stage along the central axis, allowing the distance between speaker and microphone to be varied while kept in alignment. Before each measurement, the positioning of the geometrical centre was checked using two laser beams to line up the bat-head cast and the speaker. The position of the rotary stage was adjusted using the x-y axis controller to set the crossing point of each horizontal and vertical laser beam, at both the interaural centre point of the bat-head cast and the middle point of the speaker.



(a)

(b)



(c)

Figure 5.2: Measurement photos of the new rig (a) measurement set-up in the anechoic chamber (b) bat-head cast with microphone (c) rotary table and x-y axis positioning table

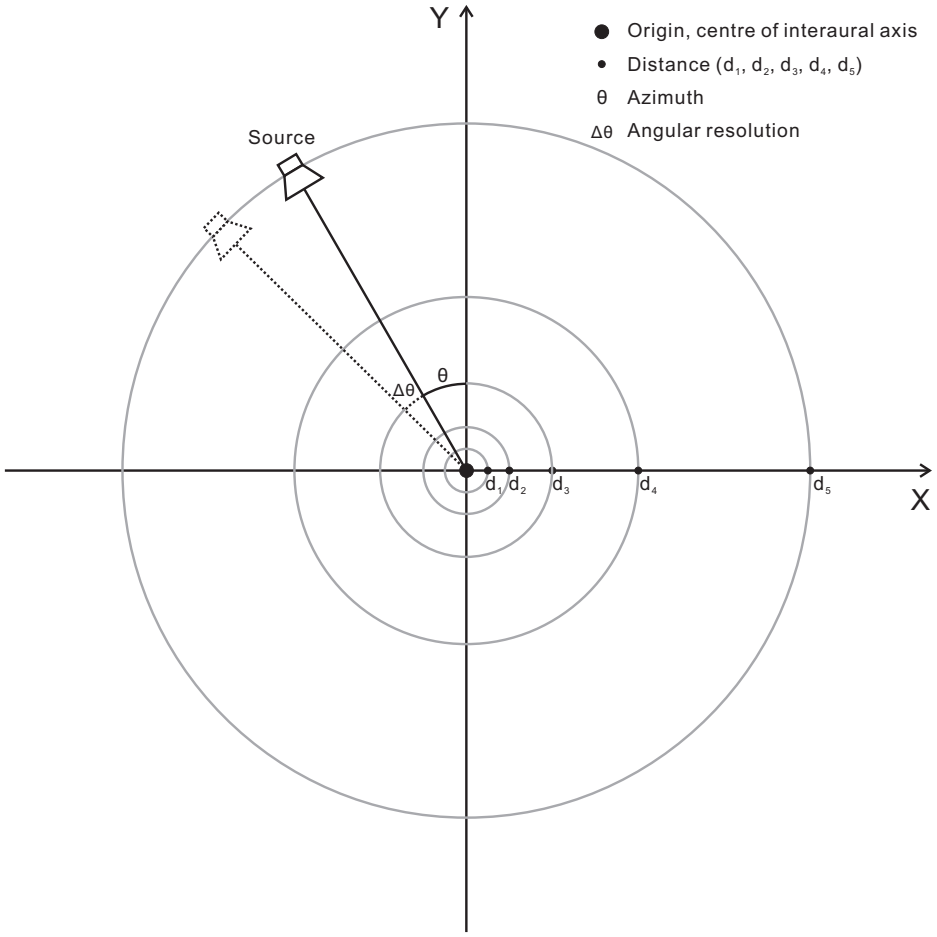


Figure 5.3: The measurement coordinates. Note that the front head is located towards positive (+) direction of Y and θ moves towards left side of the head

5.2.3 Measurement procedure

The HRTF measurement was made separately for each ear. Figure 5.3 describes the measurement coordinates used. The bat-head cast was fixed at the origin facing a positive Y-direction and the location of the sound source, θ was varied. A complete set of measurements were made at every 5° (which is denoted as angular resolution, $\Delta\theta$) step in the azimuth, 72 positions in total, in the horizontal plane for the left ear and for the right ear. After the HRTFs of the left ear were measured for all azimuths, the microphone was completely disassembled from the cast, fixed at the position of the right ear and the measurement for all azimuths were retaken. The measurement was conducted for the distances of 0.0625, 0.125, 0.25, 0.5 and 1 m between the speaker and the bat-head cast. The distances are shown as d_1, d_2, d_3, d_4 and d_5 respectively. According to the definitions of proximal and distal region, which were described in Chapter 4, the distances of 0.0625 and 0.125 m are categorised as proximal distances

whereas the distances of 0.25, 0.5 and 1 m are categorised as distal. For each measurement, a pink noise signal with 256,000 samples was generated. Then the received signal was deconvolved from the original signal to produce the impulse response. The impulse responses obtained were averaged over 16 measurements to improve the quality of recorded data by reducing the measurement errors in the amplitudes of impulse responses. Both recording and processing of the signals were conducted within the Matlab 7.2 environment. Also a pilot measurement was made at each distance in order to measure the time delay between two ears for the sound source at 0 °. Finer adjustment of the centre positioning was performed until the delay between impulse responses at each ear reduced to within a range of 4 sample. This delay corresponds to 8 μs which is approximately 5 percent of maximum ITDs expected from the measurement. The gain set-up in the Matlab program, as well as the gain of the pre-amplifier of the speakers, were adjusted until the maximum SNR of the measured impulse response was achieved. Accordingly, the gain of the conditioning amplifier connected to the microphone for the signal reception was adjusted so as not to be overloaded by the amplitude of received signals.

5.2.4 Signal-to-Noise-Ratio (SNR)

Figure 5.4 shows the impulse response at the closest (0.0625 m) and furthest (1 m) distance and their magnitudes in dB, for the measurement at 0 °. The results are shown for both left and right ears. The magnitudes were calculated by evaluating twenty times the base-10 logarithm of the squares of the amplitude (i.e $20log_{10}(\mid amplitude \mid)$) in the same time domain. The peak level of the magnitude and the noise floor level were compared to evaluate the SNR. It is considered to be important to set the gains of the amplifiers of both speaker and microphone in order to accommodate the measured response at all distances. These gains were adjusted manually until the measured data achieved a suitable SNR. The suitable SNR was determined based on the literature that a bat's ILD can reach up to 40 - 50 dB (Chapter 2). The measured SNR are shown to be approximately 60 dB for both distances. It is noted that the noise level is higher at 0.0625 m than at 1 m by approximately 5 dB. This is due to the noise generated by the speaker introducing a higher level of the signal received by the microphone as the distance decreases.

5.2.5 Interaural axis adjustment

As shown in Figure 5.5, the centre positioning of the interaural axis has been performed using two laser beams projected on the hard bat-head cast and the speaker. To do so,

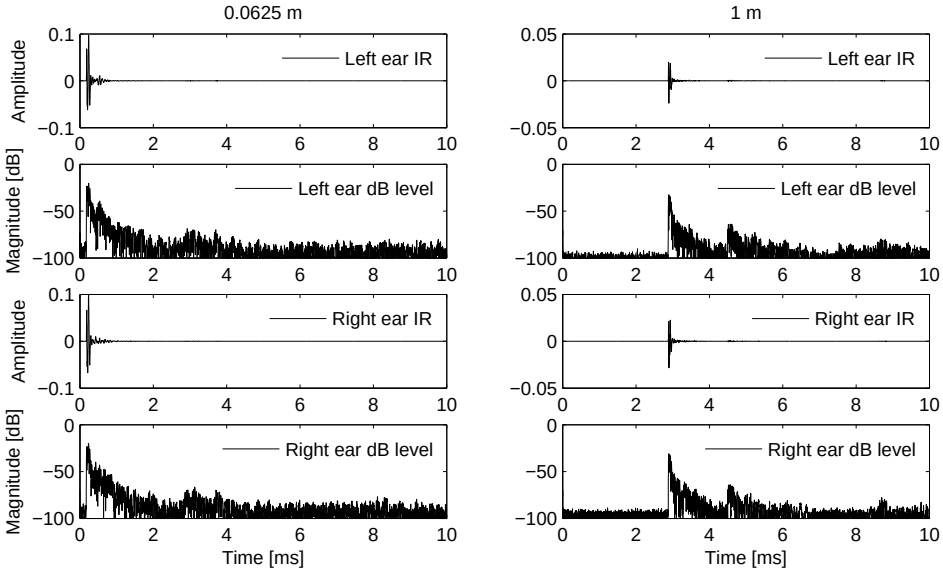


Figure 5.4: The impulse responses measured at the closest and furthest distances for the sound source located at 0 ° and SNR evaluation, 0.0625 m (left column) and 1 m (right column). Note that the amplitude range is ±0.1 for 0.0625 m and ±0.05 for 1 m respectively. The delay shown in the impulse response measured at 1 m is caused by sound travelling in the longer distance.

the bat-head cast was positioned at 0,° to face the speaker in front, then the two laser beams were projected from the back of the cast. The vertical laser beam was set to cross both interaural centre of the cast and the centre of the speaker. The horizontal laser beam was set to cross at the centre of both ears and the centre of the speaker. As the new rig is set in a fixed metal beam and the angular position of the bat-head cast is controlled by the rotary table, the measurement error is expected to be reduced compared to the first rig used in Chapter 4. The repeatability of the rig measurements will be discussed in the following section.

5.2.6 Near- and far-field examination and calibration

The free-field measurement was made using the same microphone without the bat-head cast in place. The free-field responses were measured at 5 different distances. It is considered to be important to examine the data at all 5 distances to understand the quality of the measured data from distal to proximal region. As shown in Figure 5.6, the far-field condition, which the amplitude of received signal reduces by 6 dB as the distance is doubled, is generally satisfied up to 0.125 m over the frequency range. However, the measured response at 0.0625 m only satisfies the far-field condition up to 20 kHz. As the distance between the source and receiver decreases, distortion of the received signal

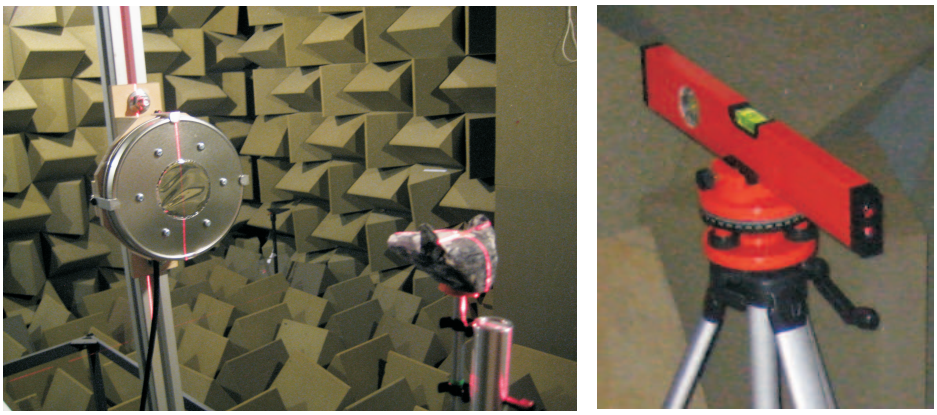


Figure 5.5: Interaural axis adjustment using the laser beams. The adjustment process (left) and the laser (right).

is expected, particularly at higher frequencies. On the other hand, the low frequency response (below 10 kHz) is shown to mostly satisfy the far-field condition across all distances. In general, the effective frequency range is considered to be from 5 kHz to 100 kHz.

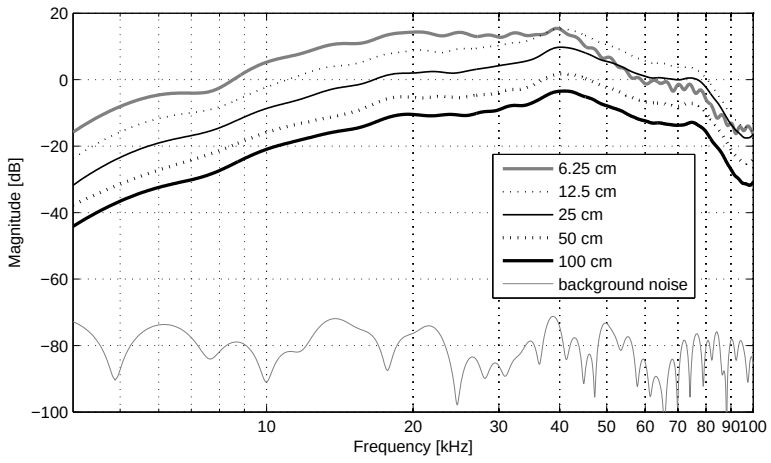


Figure 5.6: Near-field and far-field examination in all distances. Note that the background noise level is less than the level of the measured signals at the furthest distance by 40 - 60 dB across all frequencies.

The calibration of the measured HRTF data was made by using the free-field response measured at 1 m as the reference distance. This calibration was applied to all measured data at different distances to obtain the magnitude attenuation as the distance decreases. The procedure assumes that the far field condition is satisfied for all distances. However, this condition is obviously not satisfied for the measurement at 0.0625

m as shown in Figure 5.6, hence a distortion from the calibration is expected in the frequency region above 20 kHz. As the purpose of this study is to investigate the range-dependent characteristics especially at low frequency, the distortion over this particular frequency region at the closest distance are considered to be acceptable to proceed with the analysis.

To analyse the magnitude response or the interaural level differences (ILDs), all impulse responses were processed with a 250-point rectangular window around the maximum peak response and the windowed signal was zero-padded in the tail with 7942-samples to produce the post-processed 8192 samples. Then a 8192-point FFT was applied which corresponds to the frequency resolution of approximately 61 Hz in the analysis. The frequency analysis was limited to between 5 kHz and 100 kHz. The same processing was applied to the free-field response and the magnitude of the HRTF at each location was obtained by dividing the amplitude spectrum at that location by the amplitude spectrum of the free-field response. The interaural time differences (ITDs) were calculated from the cross-correlation of the envelopes obtained from the two ear measurements using the Hilbert transform in the time domain. The free-field equalisation was not applied in the processing of the ITDs to exclude the inevitable errors being induced by the centre positioning procedure as the sensitivity of the positioning error can be noticeable in the ITDs. Note that further details of the process shown above is described in Chapter 4.

5.3 Validation of the measurement system

Data validation is carried out for the new experimental set-up. The same measurement method which was described in previous section was used. Both reproducibility and repeatability of the conducted measurements were examined for each ear.

5.3.1 Reproducibility

The sound source was generated repeatedly 40 times for both ears at 90° azimuth at a fixed distance of 0.5 m to examine the performance of the speaker and microphone. The 90° angle was chosen as a representative azimuth, setting the microphone to directly face the speaker. The magnitude and the phase response of the impulse responses measured at each ear are shown in Figure 5.7. The reproducibility error was evaluated for each ear by calculating the standard deviation of the measured data for the magnitude and phase response respectively. It is shown that the magnitude error is within the variation range of less than 1 dB and the phase error is within the variation range of less

than 0.5° for all frequency regions for both ears as shown in Figure 5.7. The response measured at the right ear appeared to have less variation than at the left ear, however this difference is considered to be negligible as the measured response is expected to reach up to 40 - 50 dB. It is also interesting to note that the frequency responses at both ears are similar in the low frequency region below 20 kHz. However, each ear has different characteristics above 20 kHz showing the different notch patterns at around 80 kHz. This is considered to be due to the asymmetry between the positioning angle of the pinna at each ear. Also the pinna shape of each ear and head shape may affect the high frequency characteristics.

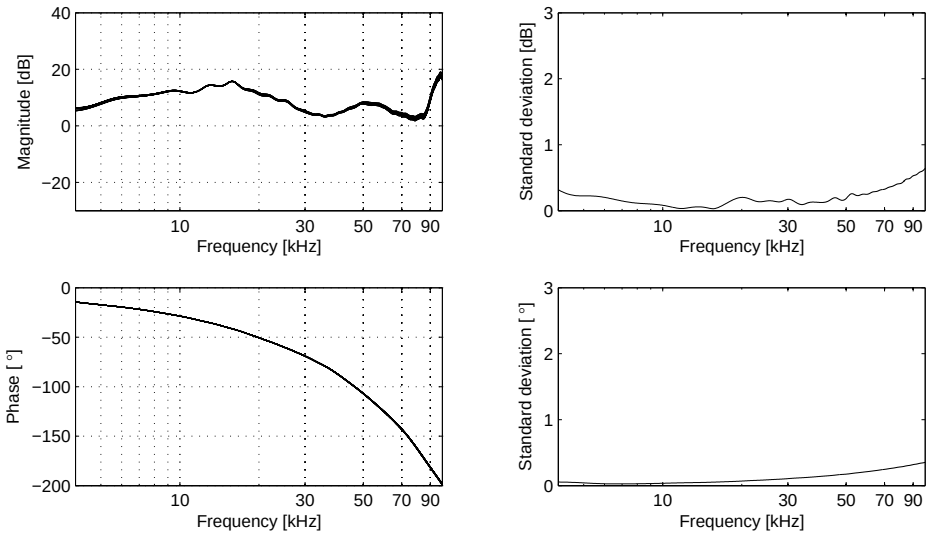
5.3.2 Repeatability

The repeatability of the procedure of inserting the microphone into the ears was also examined. The distance between the sound source and receiver was set to 0.5 m. Each measurement was made at the positions of 0°, 90°, 180° and 270°, and then the microphone and bat-head cast were completely disassembled. In the following reassembly, they were fixed in the rig again and the measurement was repeated. This procedure was repeated 10 times and both ears were examined. Based on the spectral analysis results for the measured data, the standard deviation for 10 repeated sets of measurements at each frequency was evaluated in terms of both magnitude and phase. The measured frequency response are shown in Figure 5.8(a) and Figure 5.9(a), and also the standard deviation of the magnitude and the phase responses are shown in Figure 5.8(b) and Figure 5.9(b). In general, the low frequency region below 20 kHz - 30 kHz appears to have good agreement between the 10 repeated measurements showing less than 1 dB change. These characteristics were shown to be similar across different angles in the front, back and lateral positions. However, the higher variation in the high frequency region was noticeable especially where the frequency notch is located. For the left ear at 0°, a prominent notch is located around 80 kHz and the result from repeated measurement does not correspond exactly showing a larger standard deviation reaching up to around 15 dB at maximum. It is considered to be due to the fact that several notches move around the 80 kHz range for each measurement. The standard deviation of the magnitude appeared to be well below 5 dB and that of the phase appeared to be mostly 0° in all frequency regions for the azimuth at 90° at the left ear when the microphone faces towards the sound source directly. Similar characteristics were shown for the azimuth at 270° at the right ear but the variation appeared to be larger than the responses at the left ear for the frequency region above 70 kHz. The results implies that some measurement show strong notches whereas others show weaker notches in this frequency region. The characteristics at 180° azimuth at the left ear is similar to

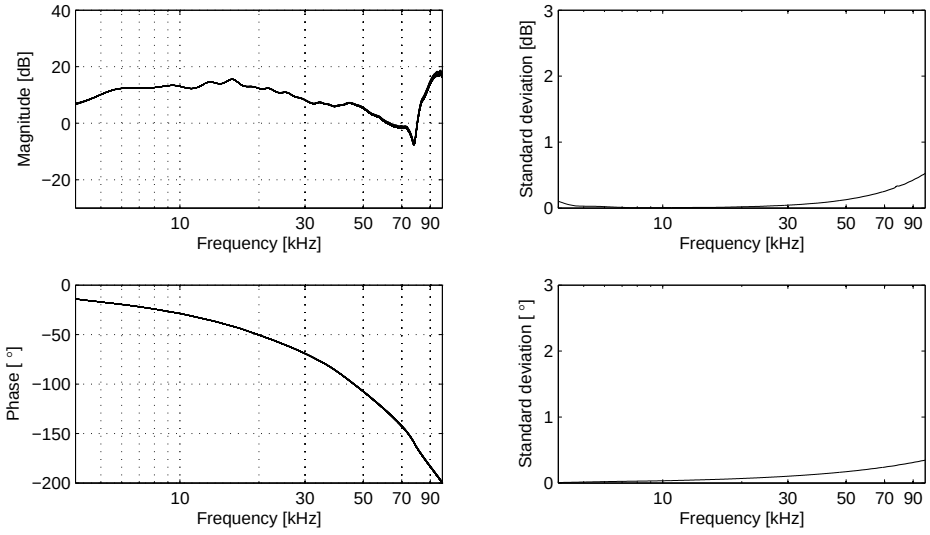
the result measured at 0°, however another notch around 40 kHz is found and the location of this notch was also shown to change around this frequency region over the 10 measurements. The sound source was positioned on the opposite side of the microphone for the left ear at the azimuth of 270°. For this azimuthal location, it is expected that there is no direct path for the acoustic sound, so the high frequency wave cannot reach the ear. The magnitude variation appeared to be larger than other source positions and the phase variation appeared to reach around 10° for the frequency region above 70 kHz. The sound source located at this azimuth does not only produce low gain but also introduces several notches. Therefore, relatively poor repeatability in this region is expected to be found. But low frequency sound can still reach the receiver as it diffracts around the head and as a result, the response shows good agreement between repeated measurements.

Similar characteristics are shown for the measurement at right ear as shown in Figure 5.9. The co-ordinate system of the azimuth is same for both left and right ears. Therefore, the response measured at 90° at the right ear corresponds to that at 270° at the left ear and the response at 270° at the right ear corresponds to that at 90° at the left ear. It was observed that the response at 270° appears to have poor repeatability above 30 kHz despite the microphone being positioned directly towards the speaker. This performance is shown to be poorer than the response at the left ear at the corresponding position. It could be due to the microphone insertion process which was conducted by hand. Considering the fact that the difference in the position of microphone at the ear is around 1 mm, this result shows how sensitive the measured response is to location.

The result not only demonstrates the higher sensitivity to the measurement in the high frequency region but also it implies a significant effect of the head and pinna on the received signals. The peak and notch characteristics in the high frequency region are mostly determined by the acoustics generated by the geometry of the head and pinna. Hence, a small variation in the position of the microphone and the positioning error of the direction of the head towards the sound source, seem to vary the physical surroundings around the microphone, hence high frequency characteristics of the received signals are affected. Therefore, it should be noted that the procedure of assembly and positioning of a sonar receiver can potentially produce unwanted peaks and notches. Especially, it is considered critical when one investigates a single peak or notch at a particular frequency appeared in the characteristics of a sonar receiver.



(a) Data measured at the left ear



(b) Data measured at the right ear

Figure 5.7: Reproducibility of the system at fixed distance. The measurement of impulse response was repeated 40 times at each ear and the head was positioned at the distance of 0.5 m. The results show the magnitude/phase responses (left column) over the frequency and the reproducibility errors were evaluated by calculating the standard deviation (right column) of the measured data. (a) left ear (b) right ear

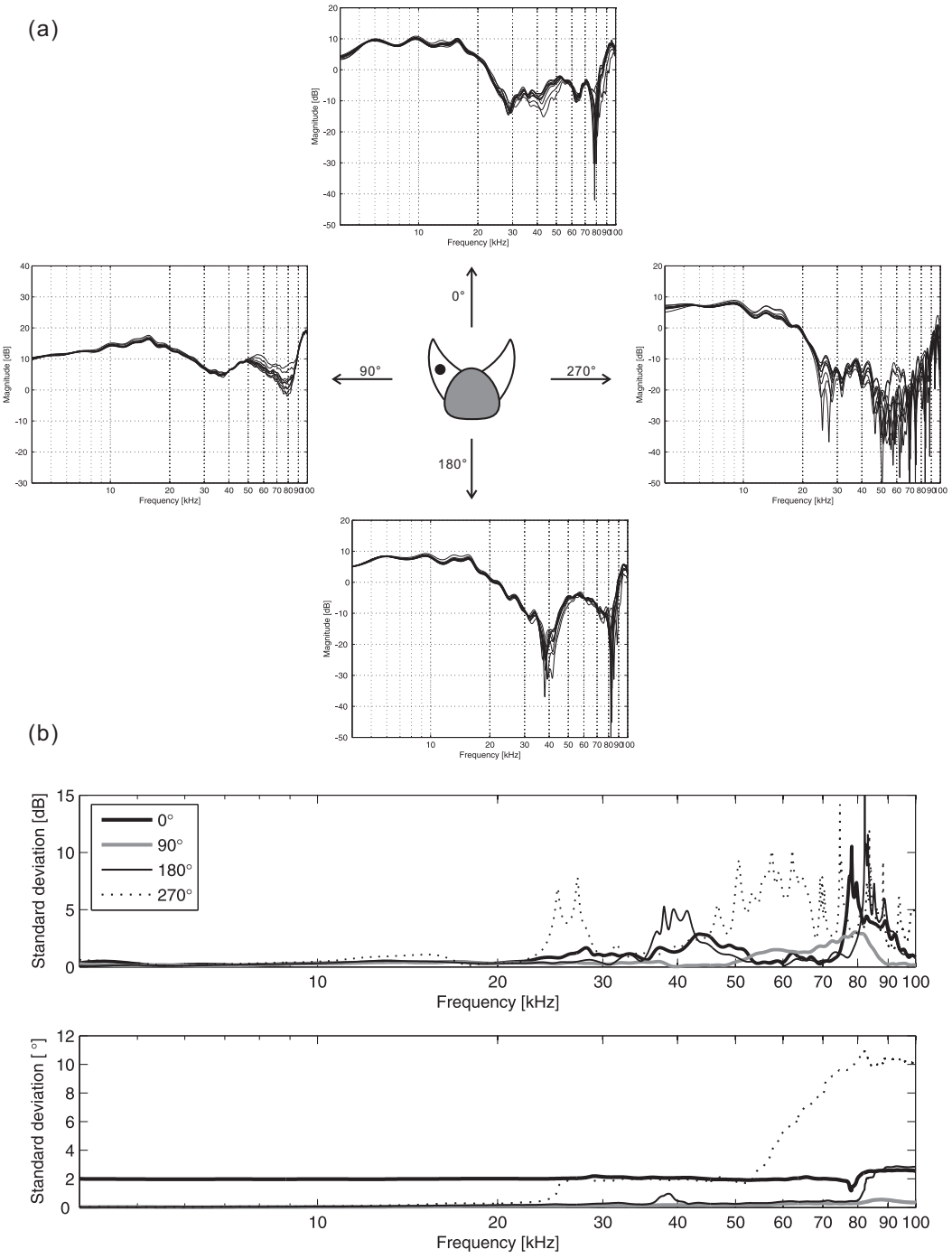


Figure 5.8: Repeatability at the left ear and the standard deviation of the measured data. (a) FFT magnitude spectrum measured at the left ear for 10 times of repetition. (b) Standard deviation at the left ear. The variation of the magnitude and the phase for each sound source positions were calculated based on the measured data shown above. The standard deviation of magnitude (top) and the phase (bottom).

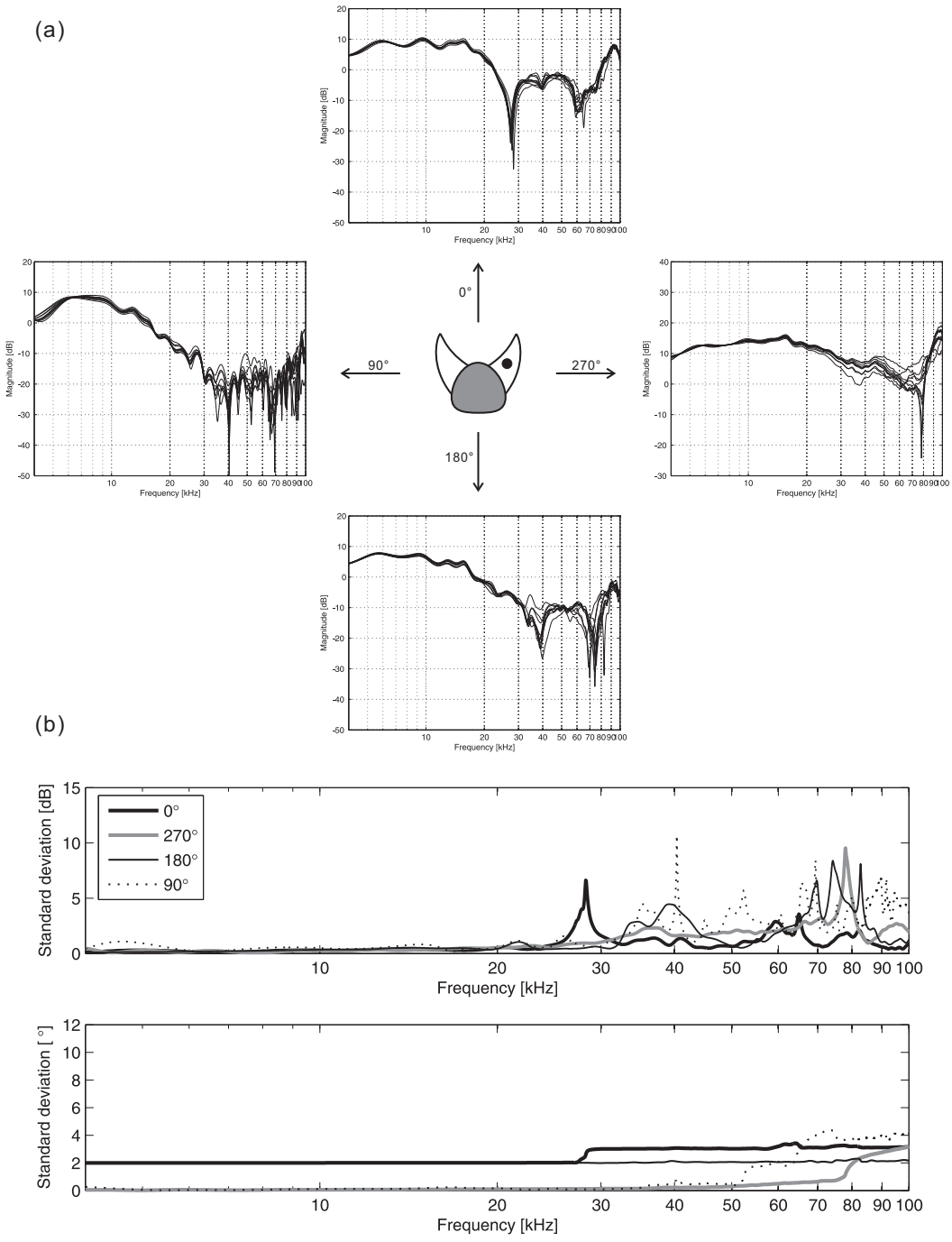
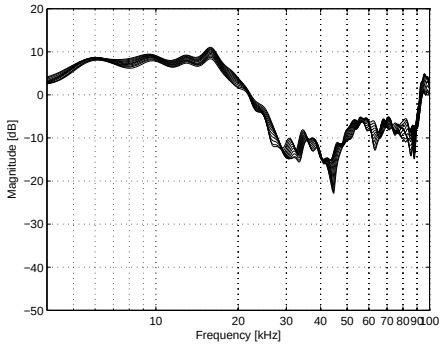


Figure 5.9: Repeatability at the right ear and the standard deviation of the measured data. (a) Frequency spectrum measured at the right ear for 10 times of repetition. (b) Standard deviation at the right ear. The variation of the magnitude and the phase for each sound source positions were calculated based on the measured data shown above. The standard deviation of magnitude (top) and the phase (bottom).

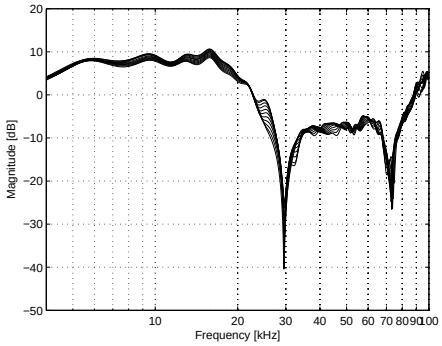
5.3.3 High angular resolution

It has been considered conventional to measure the HRTF at an angular resolution of 5° that has been considered a sufficient resolution to plot the acoustic field of the HRTF in all azimuths [Gardner and Martin, 1994]. However, the HRTF measurement of the bat-head cast in this study can introduce an error in the measured data due to a small change in the angular position. Although the rotary table used theoretically produces a 1° resolution, the manual adjustment can introduce such spatial error. Therefore, it is important to note the potential spatial error in the frequency response represented by the peak and notch structure can exist due to angular mispositioning. In this section, the result of the measured response at a higher angular resolution of 1° is described. At the 0.5 m distance, the measurement was made for 1° resolution around 4 representative source positions located on frontal midline (0° for both ears), interaural axis facing the ipsilateral ear (90° for left ear and 270° for right ear) and interaural axis facing the contralateral ear (270° for left ear and 90° for right ear) respectively. The response was measured from 355° to 5° for frontal sources, from 85° to 95° for ipsilateral sources and from 265° to 275° for contralateral sources.

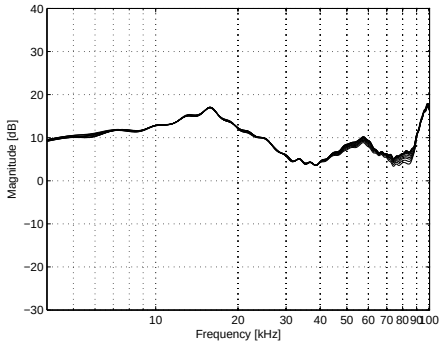
Figure 5.10 shows the measured data at each ear for the given sound sources. The results at the frontal sources of 10° variation shows that the magnitude of the frequency responses change by 1 - 2 dB below 20 kHz and around 4 - 8 dB above 20 kHz. The change is particularly prominent around the notches in the high frequency region. On the other hand, the results from ipsilateral and contralateral sounds sources varied by less than 1 dB below 20 kHz, however, the response in higher frequency region showed up to 15 dB change in peaks and notches. Therefore, ipsilateral and contralateral sources have less variation in the low frequency response compared to the frontal sources, whereas they have more variation in the high frequency response. It can be concluded that the low frequency response is more sensitive to change in azimuth when the sound source is located in frontal positions, while the high frequency response is more sensitive to the position of lateral sources.



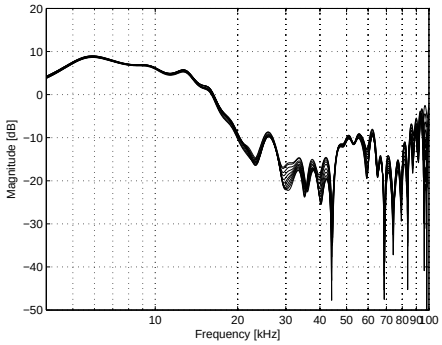
(a) 0 °, left ear



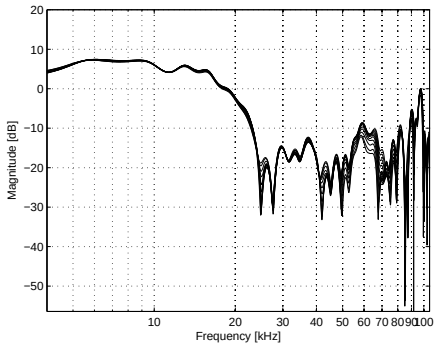
(b) 0 °, right ear



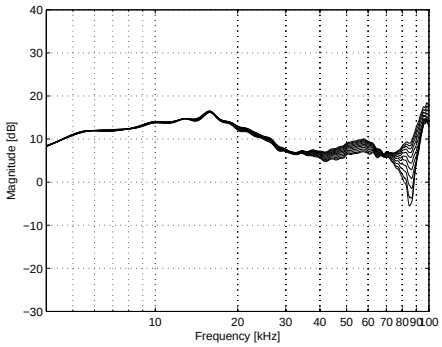
(c) 90 °, left ear



(d) 90 °, right ear



(e) 270 °, left ear



(f) 270 °, right ear

Figure 5.10: Measurement result with 1 ° resolution. Frontal sources (a) left (b) right, ipsilateral sources (c) left (d) right and contralateral sources (e) left (f) right

5.4 Monaural analysis

The monaural analysis includes the contour plots of the measured HRTF at each ear for various distances, average gain obtained from far distance to close distance and the monaural frequency characteristics depending on the location of the sound source and distances.

5.4.1 Monaural contour plots of HRTF

The contour of the measured HRTF at each ear is plotted for 5 different distances in Figure 5.11. The analysis and the representation of the data is largely based on the paper (JASA) by Brungart and Rabinowitz [1999], which describes the HRTF measurement of human model at various distances. It is also advantageous to compare the current results from the bat-cast model to those of the human model. The spectral contour plots of the HRTFs measured in all azimuths provide several characteristics explaining the acoustical phenomena around the bat-head cast. Ten contour plots at various distances from each ear are shown in Figure 5.11. The 8192-point FFT was applied to the data and the results are presented in dB at a frequency range from 5 kHz to 100 kHz. The obtained data with 5° azimuth resolution are interpolated to represent smooth contours. All contours are plotted over the range of 60 dB, ranging from -30 dB to +30 dB as shown in the colourbar.

The general characteristics of the measured HRTF in all azimuths are shown to be consistent in the 10 sets of the measurements. They are observed to have similar patterns describing the principles of the acoustics around the head. Although there exists asymmetry between the left and right halves of the cast, the general characteristics of the contour agree with each other throughout all distances. There are four major features shown in the contour plots denoted as A, B, C and D in each figure at the same locations. These areas have already been discussed in Chapter 4 in terms of general characteristics but they are discussed here in terms of the effect which is caused by the distance.

The magnitude of monaural HRTF generally increases as the sound source approaches the head and when there is a direct path between the source and ear. The high frequency doubling effect (A) is shown to be more prominent at close distance, particularly at 0.125 m and 0.0625 m. In general, the overall gain of the low frequency increases as the distance decreases and this effect is shown to be clear for both front and back

sources. The head shadowing effect (B) is observed for all distances and the multiple ridges are present in these regions. The multiple ridges in the shadowed area are introduced by the ripples produced by the wave propagation at various directions around the head and its subsequent interference. The ripples were shown to be less obvious in closer distances due to the smearing effect of acoustical sound generated from the loudspeaker. A point source, which radiates sound equally in all directions regardless of frequency, is assumed in longer distances, however this criteria is expected to change when the distance decreases due to the dimension of loudspeaker. On the other hand, it is noted that these ripples are not evident in the measurements using the soft bat-head cast but they are more prominent in the results from the hard bat-head cast. It is considered that the rigidity of the hard bat-head cast enables stronger production of acoustical reflection and diffraction. It is also observed that the shadowing of the high frequency sound first occurs at around 45° (the top of (B) area) and 135° (the bottom of (D) area). The larger shadowing area, which particularly extends to the frontal region might imply the higher directivity of the pinna.

An acoustical bright spot (C) at around 20 kHz in the low frequency region appears at all distances but the contrast between the acoustical bright spot and its surrounding area, are more obvious at the distances of 1 m and 0.5 m. The notch interference (D) is shown to be similar in the magnitude patterns between all distances at around 40 kHz - 80 kHz but the area is somewhat larger at 1 m and 0.5 m.

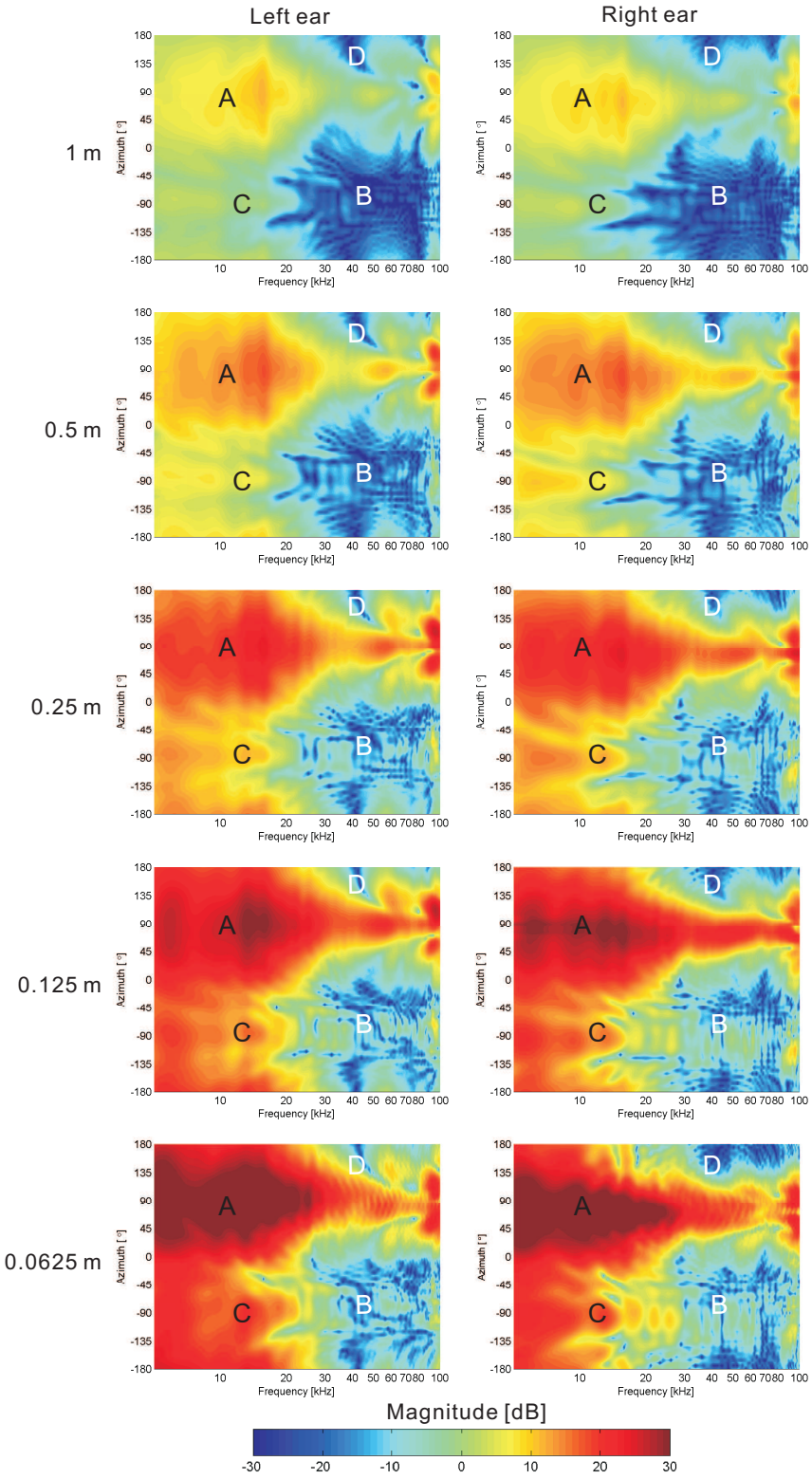


Figure 5.11: Contour plots of HRTFs in 5 distances from 0.0625 m to 1 m. The analysis method is same as shown in Figure 4.16 in Chapter 4. The magnitudes of monaural HRTFs at all azimuths are plotted over the frequency for each ear. The current figure is interpolated using the data from 5 ° step. Four reference points (labeled A-D) are present on the contour plots (see text for details).

5.4.2 Average gain in the frontal axis

The average gain obtained when the sound source moves from the furthest distance (1 m) to the closest distance (0.0625 m), was calculated for the 37 sound sources from -90° to 90° at frontal axis. The monaural frequency responses measured at 1 m were subtracted from corresponding responses measured at 0.0625 m for frontal sound sources and averaged across the azimuths. The average gain at each ear is denoted as M_L and M_R. The standard deviation across the spectral magnitudes measured at the frontal azimuths was also calculated, and denoted as σ_L and σ_R for the left and right ear respectively. The frequency range for the data evaluation was limited to 5 kHz -100 kHz. Figure 5.12 plots each average gain at each ear and the variation of $M_L \pm \sigma_L$ and $M_R \pm \sigma_R$. The minimum values of σ_L and σ_R were evaluated as 1.9 dB and 2.6 dB respectively in the low frequency region, whereas the maximum values of σ_L and σ_R were both shown to be 11.7 dB in the high frequency region. In other words, the low frequencies below 10 kHz show smaller variation across the angles but the high frequencies show relatively larger variation which is considered to be due to the peak and notch characteristics at high frequencies. Overall, more gain is obtained at low frequencies which is approximately 20 - 25 dB, than at high frequencies which approximately drops down to around 10 dB. Therefore, the low frequencies obtain relatively larger gain than the high frequencies as the sound source approaches the bat-head cast. This result implies that the characteristics in the low frequency range might provide a role in distance perception. The definition of low frequency might vary depending on the size of the head as the low frequency diffraction and high frequency reflection occur, based on the relationship between the dimensions of the head and the wavelength at each frequency. Duda and Martens [1998] has adopted a normalised distance which describes the ratio between radius of a sphere (head) and distance from source to the centre of the sphere, and has shown the range-dependence of HRTFs with frequency. Brungart and Rabinowitz [1999] have shown that the low frequency effect is observed in human sound localisation when the HRTFs of KEMAR are measured at different distances. It is clear that this effect is an acoustical phenomenon based on the physical properties of frequency, head-size and distance. Hence, the frequency characteristics of the HRTFs can be theoretically calculated and expected if the distance and the head-size are known. In the case of the bat-head cast in this study, the contrast of the gain between low and high frequencies is as large as 10 dB. This implies that a larger gain benefits better SNR in the echolocating environment where the echo intensity reduces to 10 - 20 dB (as reviewed in Chapter 2), hence the low frequency information is potentially considered to be useful in the processing of returning echoes.

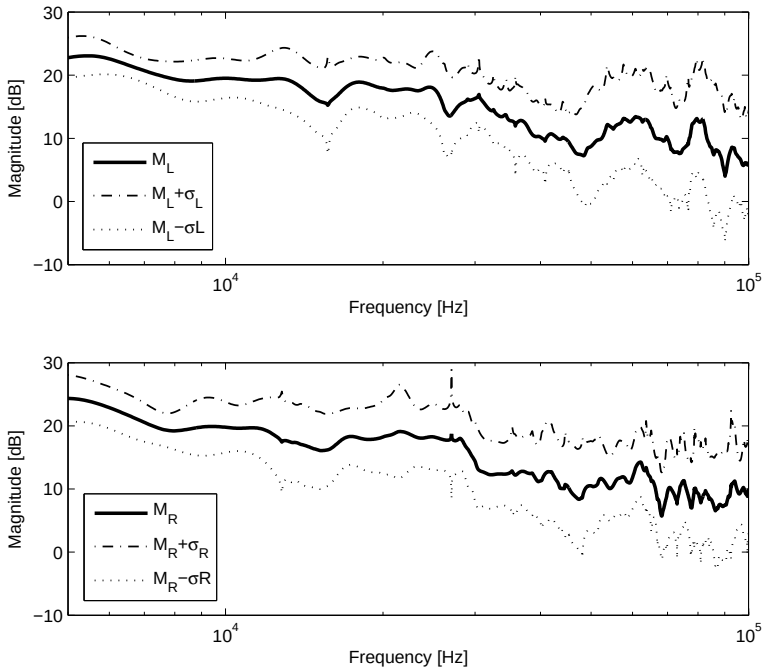


Figure 5.12: Average gains (M_L, M_R) obtained from 1 m to 0.0625 m, left ear (top) and right ear (bottom). See text for details.

5.4.3 Distance effect and source positions

For further investigation of the distance effect on the frequency response, the frequency responses at source positions of $-90\,^\circ$, $-60\,^\circ$, $-30\,^\circ$, $0\,^\circ$, $30\,^\circ$, $60\,^\circ$ and $90\,^\circ$are shown in Figure 5.13 for 5 chosen distances. The monaural HRTF in each source position is calculated by subtracting the magnitude of the free-field frequency response from that of the measured HRTF. Depending on the source positions, the characteristics of the distance effect also change. It is noted that, for $-30\,^\circ$, $0\,^\circ$ and $30\,^\circ$, the gain in the high frequency region does not appear to have substantial change although the distance varies but it shows more variable characteristics, whereas the gain in the low frequency region increases systematically as the source approaches the bat-head cast. On the other hand, this systematic change in low frequencies below 10 kHz expands towards the high frequency region for the ipsilateral sources at $60\,^\circ$ and $90\,^\circ$ showing linear increase in gain as the distance decreases. The low frequency diffraction around the head results in the systematical change in the gain when the sound source is located at the angles close to the midline in frontal axis. On the other hand, the high frequency signal only reaches at the ears when there exists direct sound travelling path. Note that the linear increase in gain is not obvious for the distance of 0.0625 m due to the limitation of the measurement in near field and its frequency response.

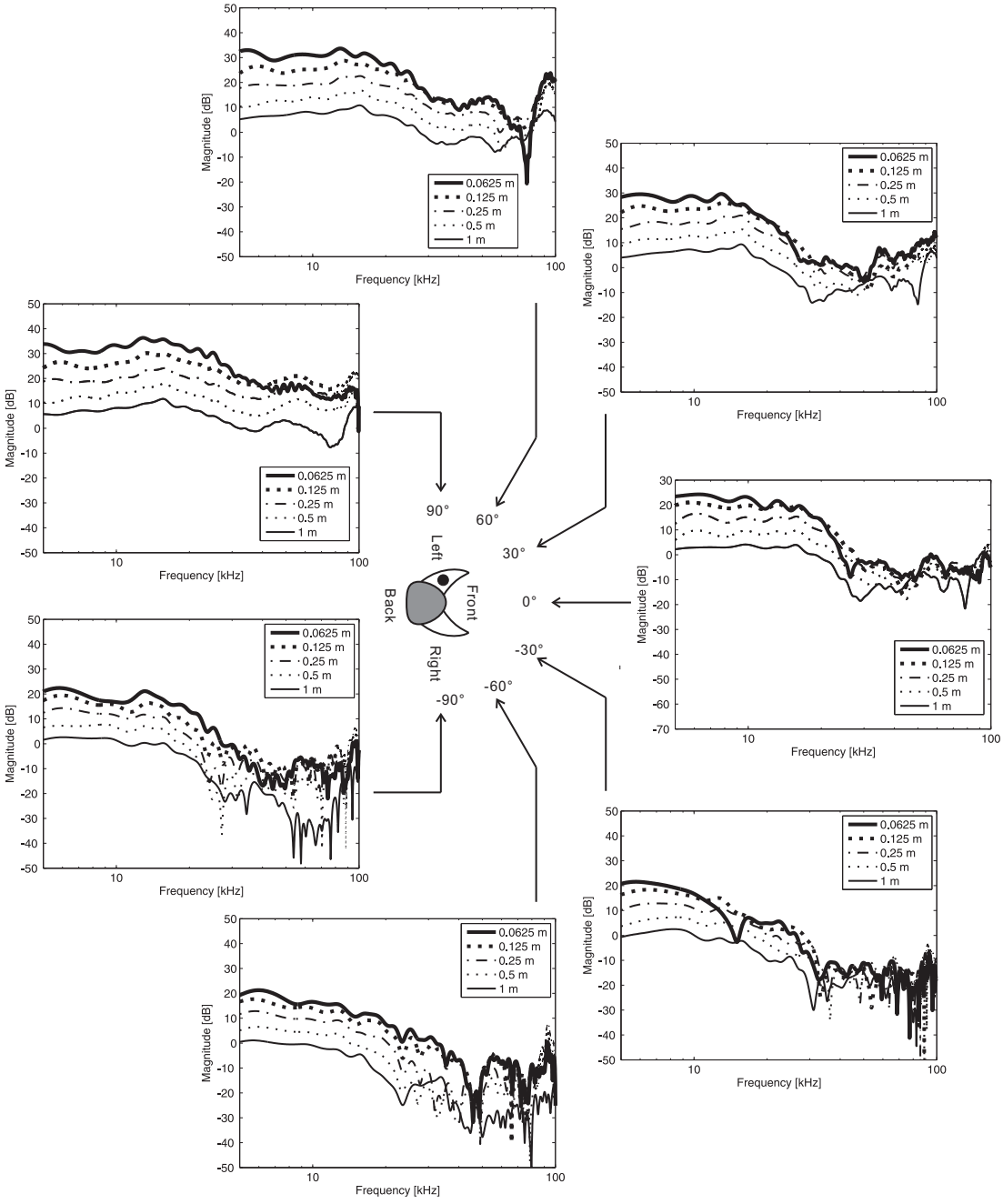


Figure 5.13: Monaural characteristics of the measured HRTFs showing the distance effect depending on source positions of −90 °, −60 °, −30 °, 0 °, 30 °, 60 ° and 90 ° using the position of the left ear as a reference

5.4.4 Summary of monaural findings

(1) The monaural characteristics of the measured HRTFs were shown to be similar across various distances from 0.0625 m to 1 m in terms of acoustical pattern in the contours for all azimuths around the head. However, the magnitude increases more rapidly at the low frequency (around 10 kHz) than at the high frequency. The average gain obtained between the furthest and the closest distance appears to be approximately 10 dB larger at the low frequencies than at the high frequency. Therefore, the sound source reaching each ear is effectively low-pass filtered and this phenomenon is shown to be similar to that in human HRTFs.

(2) The linear increase in the gain as the sound source approaches the bat-head cast is found in the low frequency regions across all sound sources located in frontal azimuths. However, this effect is limited to the high frequency region only when the sound sources are located near to the ipsilateral ear.

5.5 Binaural analysis

Here, binaural analysis is carried out based on the interaural differences of the measured response at both ears. The interaural level difference (ILD) analysis is made in relation to both azimuths and distances, then the acoustic field of ILDs is analysed. Also the distance effect on the ILD change is investigated in different frequency regions. The concept of angular sensitivity of the ILD change in the frontal axis is used to investigate the relationship between binaural hearing, distance and frequency. Finally, interaural time differences (ITDs) were calculated for 5 distances and plotted over all azimuths.

5.5.1 Acoustic field of ILDs

ILDs indicate the magnitude difference between two ears when a sound is generated from a particular angle and location. In this study, the ILDs were calculated by subtracting the magnitudes between the responses at left and right ears using the 1/3 octave band frequency analysis from 8 kHz to 100 kHz which consists of 12 different frequency bands. The centre frequency of the octave band analysis being applied is shown in Table 5.1. The centre frequency is set to have same interval in log scale and the bandwidth for each frequency was determined based on 1/3 octave band calculation for upper and lower frequencies. Therefore, the bandwidth increases as the frequency increases. For example, the lowest centre frequency (8 kHz) has a bandwidth of 1839 Hz whereas the highest centre frequency (100 kHz) has a bandwidth of 22156 Hz. The octave band analysis plots the ear response in a similar manner to the hearing

mechanism and it suppresses unwanted noise in the high frequency response as the wider bandwidth contains relatively more energy of the signal compared to the noise. Figure 5.14 plots the frequency response of 1/3 octave bandpass filters.

Table 5.1: 1/3 octave band analysis. For convenience, f_1, f_2, f_5, f_8, f_9 and f_{12} are rounded up to 8 kHz, 10 kHz, 20 kHz, 40 kHz, 80 kHz and 100 kHz respectively.

Label	f_c [Hz]	f_l [Hz]	f_h [Hz]	Bandwidth [Hz]
f_1	7943 (8 kHz)	7077	8916	1839
f_2	10000 (10 kHz)	8909	11225	2316
f_3	12589	11216	14131	2915
f_4	15849	14120	17790	3670
f_5	19953 (20 kHz)	17776	22396	4620
f_6	25119	22378	28195	5817
f_7	31623	28173	35495	7323
f_8	39811 (40 kHz)	35467	44686	9219
f_9	50119	44651	56256	11606
f_{10}	63096	56212	70823	14611
f_{11}	79433 (80 kHz)	70767	89160	18394
f_{12}	100000 (100 kHz)	89090	112246	22156

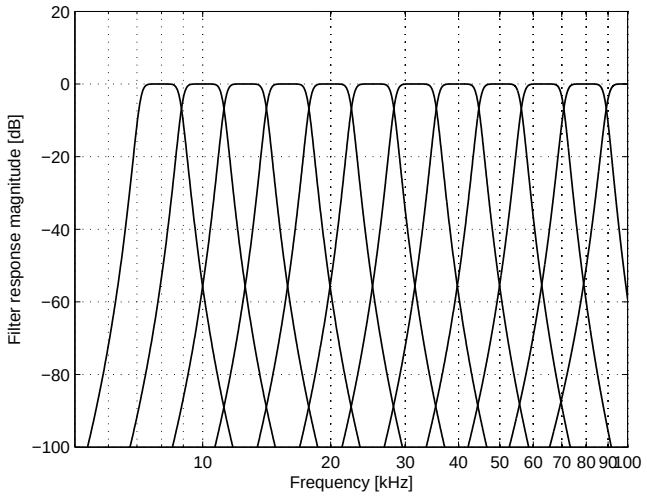


Figure 5.14: 1/3 octave bandpass filter responses being used for ILD acoustic field analysis. The centre frequency is set to have same interval in log scale.

The ILDs measured at each distance are plotted over the azimuth for the 12 frequency regions listed in Table 5.1. The low (f_1-f_4), intermediate (f_5-f_8) and high (f_9-f_{12}) frequency regions were analysed separately. Figure 5.15 shows the absolute values of measured ILDs at the distances of 1 m, 0.5 m, 0.25 m, 0.125 m and 0.0625 m. The absolute values of ILDs were evaluated to present the level difference between two ears as positive values. Note that the coordinate of the polar plot is as same as depicted in Figure 5.3. The ILDs converge to 0 dB when the source is located in midline (0° and 180°) for all frequencies. As the sound source moves towards lateral positions (90° and −90°), the absolute value of ILD increases. The two peaks around both 90° and −90° at low frequency are observed for the far distance such as 1 m and they are clear for the measurement at the left ear. Yet, they are not obviously seen in the intermediate and high frequency regions except the result of f_5 at the left ear. This is considered to be due to the low frequency diffraction which produces relatively higher gain at the ear which is located away from the sound source as the wavelength of the low frequency is long enough to allow the sound waves to travel around the head. The magnitude of the ILD in the low frequency region also increases as the distance reduces. For example, the peak value of the ILD at f_1 (8 kHz) was shown to be less than 10 dB at 1 m but it increased up to around 15 dB at 0.0625 m. However, the intermediate and high frequencies tend to maintain the gain as the distance varies. In addition, at low frequencies, the pattern of ILDs shows smooth change in values as the azimuth varies, whereas the pattern produces multiple irregular peaks and notches as the frequency increases. Furthermore, the asymmetry between the left and right halves of the cast were observed in the plots. It is interesting to note that the asymmetric characteristics are more prominent at lower frequencies than at higher frequencies. Also among the low frequencies, the difference between the left and right is larger for the ILDs measured at the closest distance. It can be concluded that the effect of geometrical difference between the left and right heads, are more sensitive at lower frequencies and at closer distances.

Based on the results shown above, the ILD acoustic field in the measured space was plotted. 8 kHz and 10 kHz for the low frequency region, 20 kHz and 40 kHz for the intermediate frequency region and 80 kHz and 100 kHz for the high frequency region were chosen for analysis and the results are shown in Figure 5.16. The results are plotted over 5 distances in the contours for those frequency regions. Each contour shows the ILD in distances from 0.625 m to 1 m for all azimuths in the horizontal plane. The magnitude values between the distances are interpolated to smooth the results. The bat-head cast is located in the centre of each plot. The magnitude of all contour plots

are fixed from −25 dB to 25 dB to show the general patterns depending on the frequency.

The ILDs are shown to converge to 0 dB at the centre of the frontal axis for all frequencies. This is due to the fact that the magnitude of the sound reaching two ears is expected to be same when the sound source is located in line with frontal-central axis. The results for 8 kHz and 10 kHz show the low frequency characteristics. In this frequency region, the ILDs are generally smaller than those at higher frequencies. It is clear that the ILDs reduce as the source approaches the central axis, however the area nearby the central axis where the ILDs are small is larger for the low frequency compared to the intermediate and high frequency regions. The acoustic bright spots are also found around 90° and −90° where the low frequency diffraction occurs and the ILDs decrease. It is notable that the ILDs increase as the source locations at the lateral sources approach towards the central origin of the contour (where the bat-head is located). The gain in the proximal region (closer than 0.125 m) is higher than in the distal region (further than 0.125 m). This phenomenon is less obvious in the intermediate and high frequencies such as 20 kHz, 40 kHz, 80 kHz and 100 kHz. It is also clear to observe that ILDs at low frequencies over the lateral sound sources in the distal region (nearby the outer circle in the plots) increase as the sound source moves away from the frontal axis at a fixed distance, however they decrease as the sound source approaches to the position in line with the ear.

The ILD characteristics of the intermediate and high frequency regions are distinguished from those of the low frequency region particularly in the lateral source positions. The largest gain is not only shown for lateral source positions but also the largest gain continuously appear in wider area of lateral azimuthal positions compared to the result shown in the low frequency regions. Also the contrast between the gain values is stronger than that in low frequencies, i.e. the difference between maximum and minimum ILDs is appeared to be larger at the higher frequencies. However, there is less systematic variation in gain for lateral sound sources compared to the results shown in the low frequencies.

This empirical result may have perceptual implications for the processing of echoes at various distances using various frequencies as follows.

- Different frequency ranges might have various roles depending on the azimuthal position and the distance of the sound source.
- Low frequency might be effective in distinguishing the direction of lateral sound sources in distal region as the gain systematically change over the lateral positions.

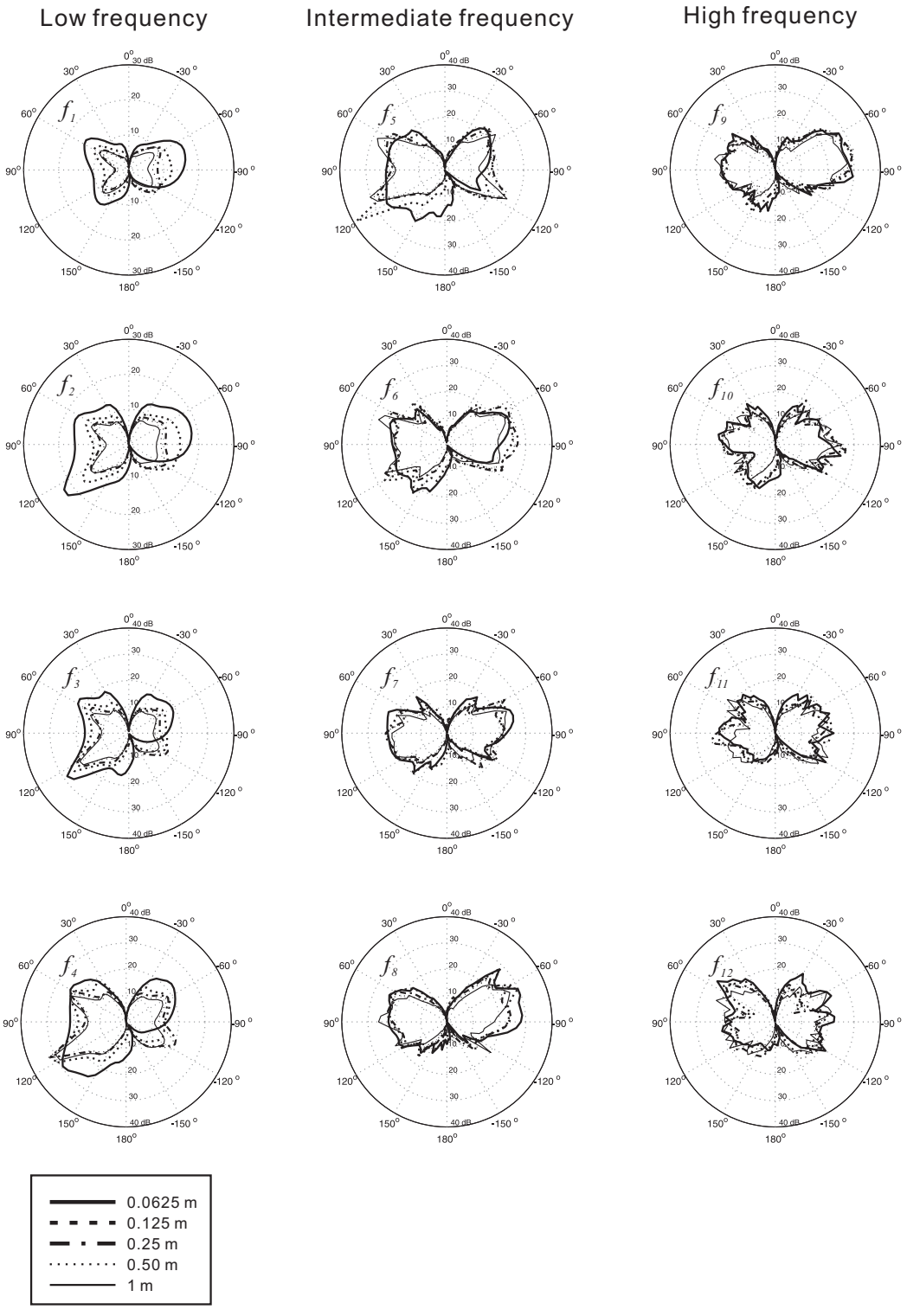


Figure 5.15: ILDs vs. azimuths over all frequency regions. The bat-head faces towards the direction of 0 ° and the source locations are positioned at each angle represented as degrees. The magnitude of ILDs are plotted as absolute value, hence the data from the right half is also represented as positive value as well as those from the left half. Note that the magnitude scales of the plots for f_1 and f_2 are 0 - 30 dB for better view whereas others are 0 - 40 dB

- High frequency might be effective in the detection of distal sound sources as it has higher gain and provides obvious differences between left and right.

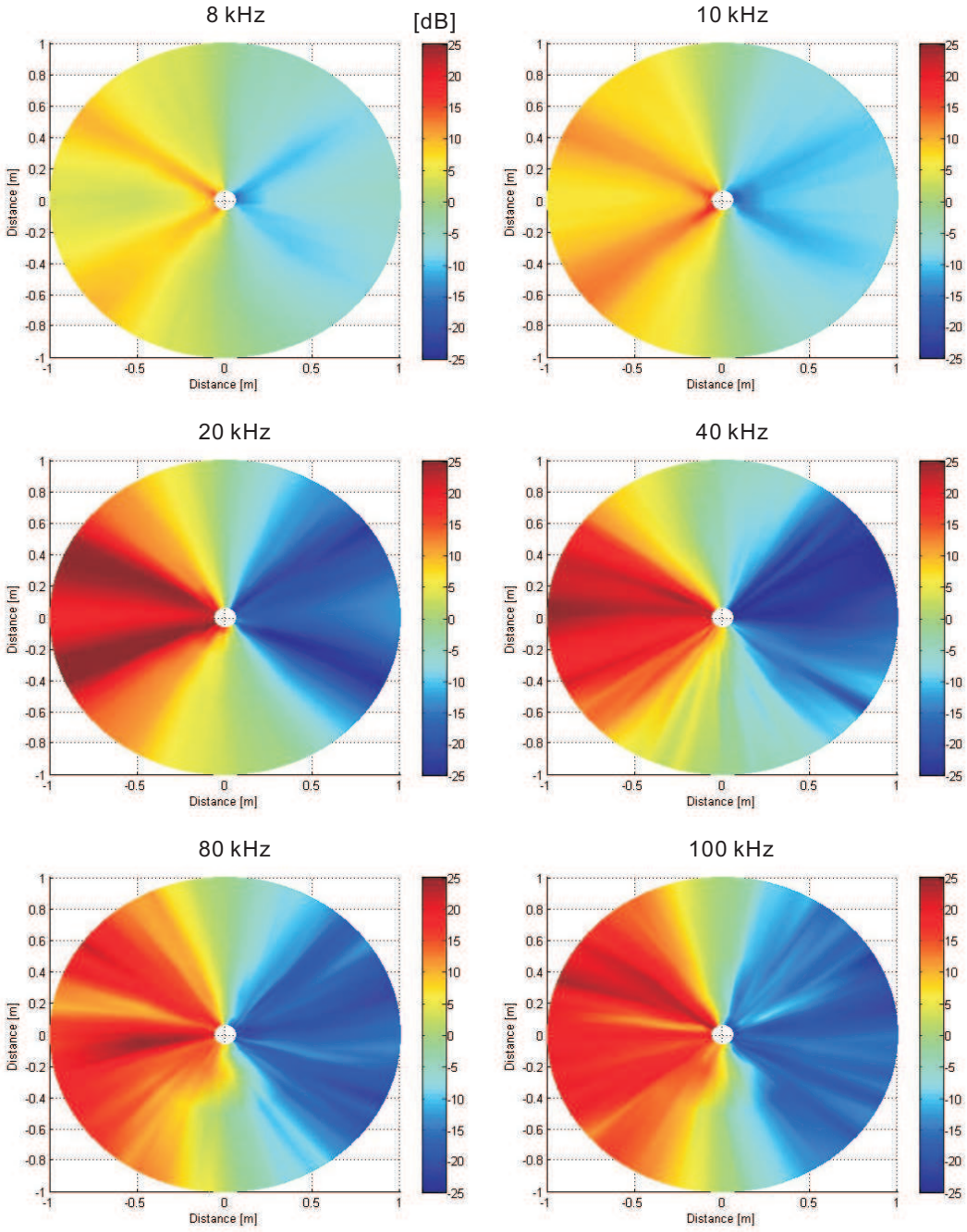


Figure 5.16: Acoustic fields of ILDs over the position of the sound source. The coordinate of the plot is as same as shown in Figure 5.3. Note that the range between the sound source and the bat-head cast is from 0.0625 m to 1 m in all azimuths from 0 ° to 360 °. The ipsilateral side is shown in the left-half of the plot and the contralateral side is shown in the right-half. The magnitude values between the distances are interpolated using the results from the measured data at 0.0625 m, 0.12 m, 0.25 m, 0.5 m and 1 m.

5.5.2 Distance vs. frequency

For each frequency region from f_1-f_{12}, the calculated ILDs are plotted in Figure 5.17 over the distance for the azimuths of 15°, 30°, 60° and 90°. The results show how the ILD changes with distance and the characteristics appear to be different depending on the frequency region. For low frequency, it is obvious in the proximal region (less than 0.25 m) that the ILD increases as the distance reduces. These characteristics are shown for the frequencies of f_1, f_2 and f_3, however, the decrease in ILD according to the distance in this region is not clear for f_4 at the azimuths of 15° and 30°. On the other hand, for the intermediate and high frequencies, the ILDs does not increase obviously in the proximal region as the distance decreases and the change is not consistent across these frequencies. It is considered to be the result from irregular peak and notch characteristics. It is interesting to note that the difference between the ILDs of f_1 measured at 15° and 30° is as small as 1 dB at far distances and this difference becomes larger in higher frequencies (f_2 - f_{12}). Therefore, it appears possible to distinguish between the source positions of 15° and 30° at higher frequencies. For the lateral source positions of 60° and 90°, the level difference between them is relatively consistent across the distances at the low frequencies, however the difference appear to vary irregularly at the intermediate and the high frequencies.

To investigate the distance effect on the binaural cues further, the ILD results at 8 kHz, 10 kHz, 20 kHz, 40 kHz, 80 kHz and 100 kHz are plotted in Figure 5.18 over the different angles in the frontal axis for 5 distances. In general, the ILDs increase as the sound source moves towards the lateral positions and this tendency is similar at low, intermediate and high frequencies. However, the ILD pattern at 8 kHz and 10 kHz changes and the maximum ILDs increases as the distance decreases whereas the ILD patterns at intermediate and high frequencies maintain similar trend at various distances. Also it is noted that the largest ILD value does not appear at 90° but the ILD tends to reach the maximum values around 60° to 90°.

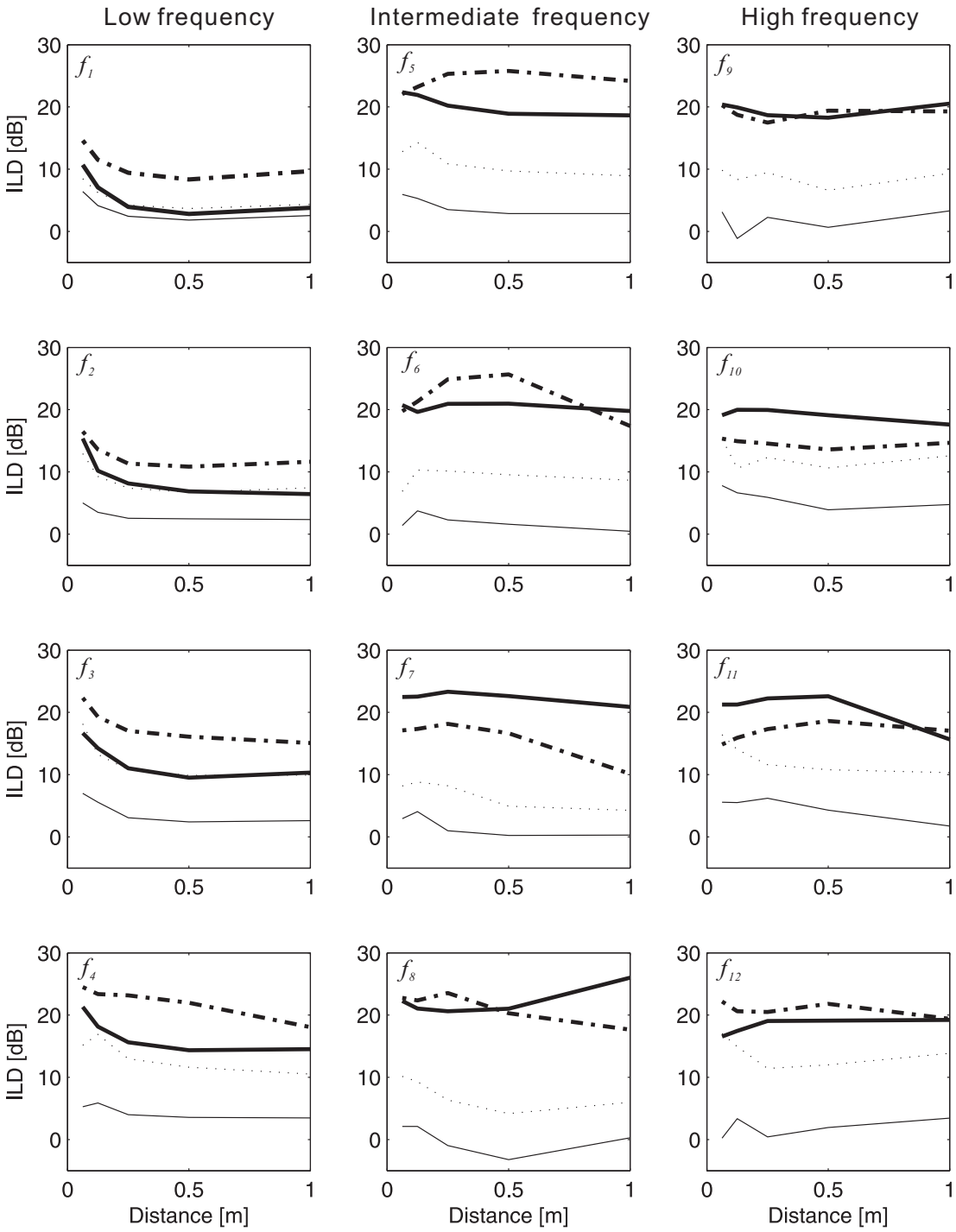


Figure 5.17: ILD vs. distance for all frequency regions from f_1 to f_{12} (15° - thin solid line, 30° - dotted line, 60° - thick dash-dot line and 90° - thick solid line).

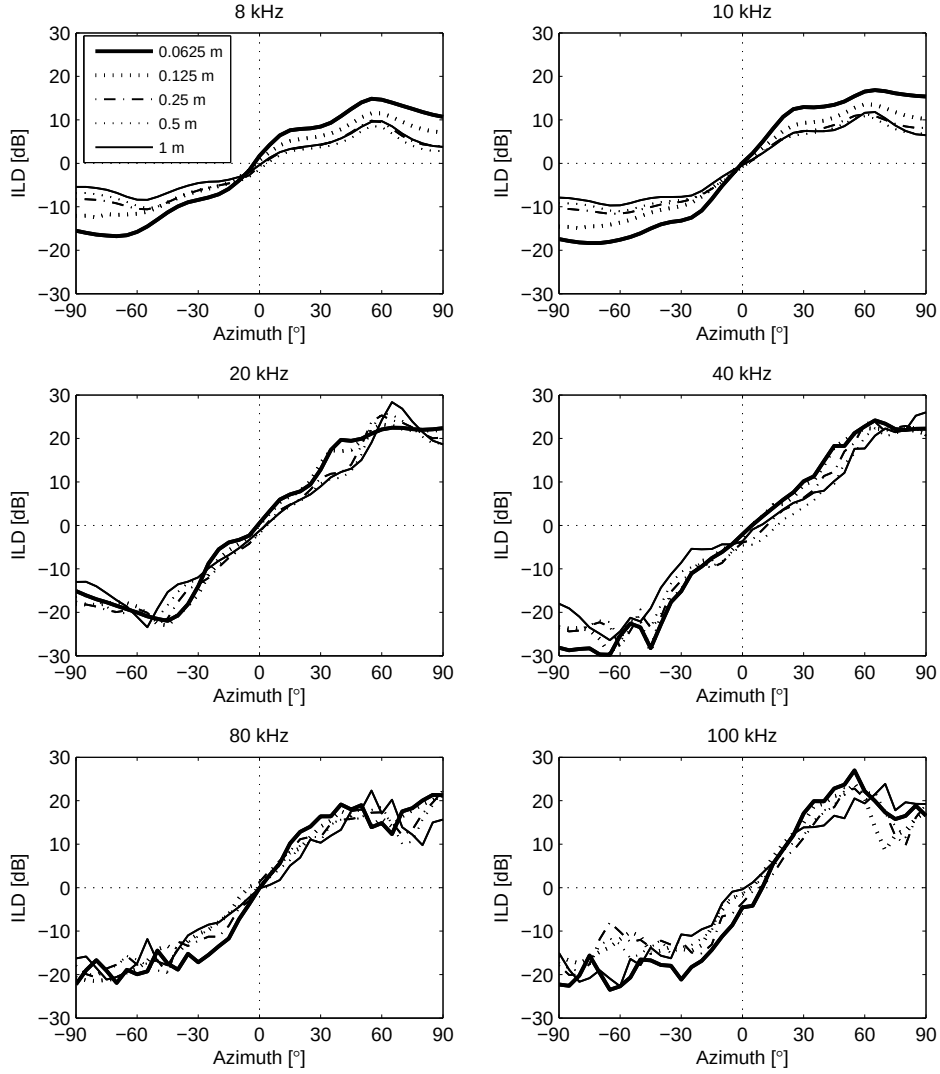


Figure 5.18: ILDs at 8, 10, 20, 40, 80 and 100 kHz in 5 distances over the frontal azimuths. The linear increase of ILDs with increase in azimuth is shown for the range approximately from $-30\,^\circ$ to $30\,^\circ$ for all frequencies. However, the ILDs in the range of $-90\,^\circ$ - $-60\,^\circ$ and $60\,^\circ$ - $90\,^\circ$ are relatively less dependent on the azimuth. The effect of distance appears clear for the frequency regions of 8 kHz and 10 kHz, but not for the higher frequency regions.

5.5.3 Comparison to sphere model results

A sphere model provides theoretical insights into the physical characteristics of ILDs. First of all, the sphere model result confirms the reliability of the measured data. Secondly, the flexibility in computational model which enables us to vary the size of head, frequency and the distance provides opportunities to theoretically understand the function of HRTF in bats which naturally change the frequency and distance during echolocation. To provide better insight into the analysis of measured HRTFs, the sphere model prediction is made in this section and compared to the measured data. This study has adopted the theoretical solution to prediction of HRTF suggested by Duda and Martens [1998] on a rigid spherical head model. This solution is based on the prediction of sound pressure level on the surface of the sphere which was originally proposed by Rabinowitz et al. [1993]. Whereas most of the previous studies about acoustical prediction in relation to reflection and diffraction have adopted Rayleigh's method, Rabinowitz *et al.* have proposed a solution to calculating the sound pressure level of a sinusoidal point source at any range of r greater than the radius of a sphere a. According to the solution provided, the free-field sound pressure (p_{ff}) at a distance r from a sound source can be written as

$$p_{ff}(r,\omega,t) = -i\omega\frac{\rho_0 S_\omega}{4\pi r}e^{i(kr-\omega t)} \tag{5.1}$$

where ρ_0 denotes the density of air (kg/m^3), S_ω is source strength magnitude of flow from an ideal point source (m^3/s) and k denotes the wave number (m^{-1}, ω divided by speed of sound c). The sound pressure on the surface of the sphere (p_s) is written as

$$p_s(r,a,\omega,\theta,t) = \frac{i\rho_0 c S_\omega}{4\pi a^2}\left(\sum_{m=0}^{\infty}(2m+1)P_m(cos\theta)\frac{h_m(kr)}{h'_m(ka)}\right)e^{-i\omega t} \tag{5.2}$$

where P_m is the Legendre polynomial function of degree m, h_m is the mth-order spherical Hankel function, and θ denotes the angle of incidence.

The head-related transfer function H can be defined as the ratio between the sound pressure in the free-field and the sound pressure which is modified on the surface of the sphere. Therefore, H is defined by

$$H = 20log_{10}\left|\frac{p_s}{p_{ff}}\right| \tag{5.3}$$

The result of this theoretical solution enables us to compare it to the measured data from the bat-head cast. The spherical model is simulated when the radius of the sphere is set to 2 cm which is the size of the bat-head cast, for the same frequency range as the measured HRTFs. Figure 5.19 shows both predicted ILDs from the sphere and the

measured ILDs from the bat-head cast for the frontal azimuth from -90° to 90° at all distances. The results are shown for the frequency region of 10 kHz, 20 kHz, 40 kHz and 100 kHz. The magnitudes of ILDs from the theoretical result ranges between approximately ± 20 dB at the furthest distance, and ± 30 dB at the closest distance. The result from the measured data shows that the ILDs range between approximately ± 25 dB at the furthest distance, and ± 30 dB at the closest distance. Considering the asymmetry of the dimension of the bat-head cast which has longer distance of front-to-back than that of side-to-side, the range of ILDs from the experimental results seem to agree with the simulated results. The ILD values generally increase as the angle of the source moves towards the ipsilateral positions. Also both results show maximum peaks off $\pm 90^\circ$ in general. It is regarded to be due to the effect of the acoustical bright spot which appears around $\pm 90^\circ$ and eventually reduces the level difference between the sounds reaching at two ears. The sphere model result also shows that the ILDs increase monotonically approximately from -60° to 60°. However, it is noted that the measured data for the frequency region of 100 kHz does not show obvious maximum peaks off $\pm 90^\circ$ at close distances such as 0.0625 m. It is considered as the effect of near-field where the frequency characteristics are distorted at high frequencies (see section 5.2.6). On the other hand, it is clear that larger differences between minimum and maximum magnitudes of the ILDs are present at the higher frequencies. Furthermore, the ILDs in higher frequencies increase more sharply as the angle moves towards the ipsilateral sources. It implies that the binaural gain at higher frequency changes more rapidly than at low frequency as the location of the sound source moves towards the frontal positions. It is also noted that the asymmetry between the left half and the right half affects the data especially measured at lateral source locations from 60° to 90°, and from -60° to -90°.

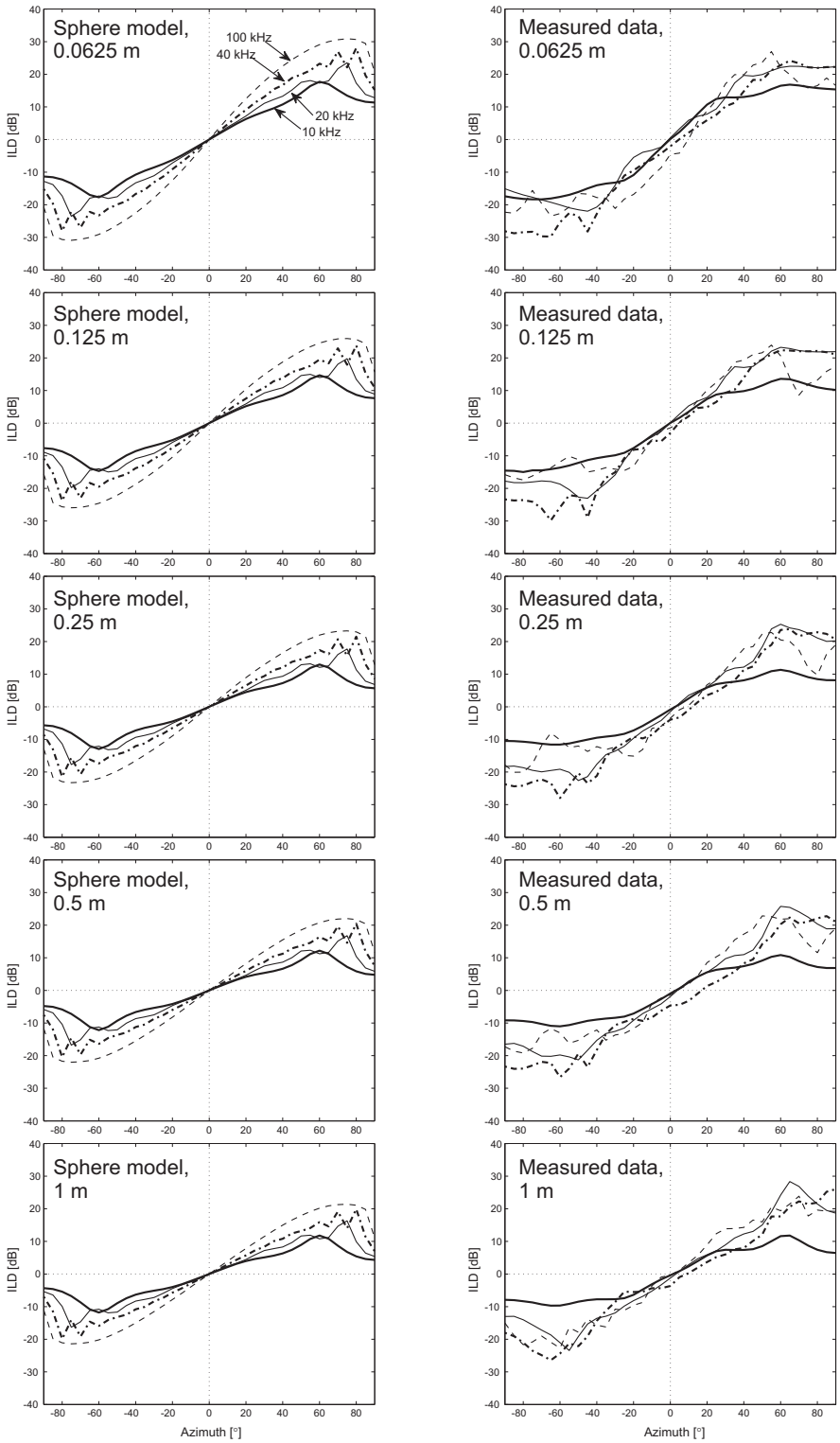


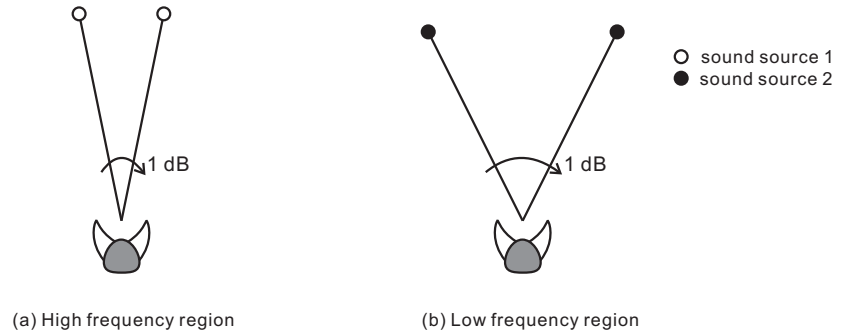
Figure 5.19: ILDs vs. azimuths for the frequency region of 10 kHz, 20 kHz, 40 kHz and 100 kHz at all distances (left column - sphere model, right column - measured data).

5.5.4 Angular sensitivity

Based on the ILD plots shown above, the angular resolution of the ILD slope is calculated. The concept of the ratio between the azimuthal angle and the ILDs has been used in previous studies to describe the performance of the binaural hearing using ILD cues [Obrist et al., 1993; Firzlaff and Schuller, 2003]. In this study, the angular resolution is obtained by extracting the ILD data between the $-30\,^\circ$ and $30\,^\circ$and calculating the first order coefficient of its slope assuming the ILD increases monotonically in this azimuthal region. The coefficient of a polynomial function of degree 1 was calculated to obtain the minimum error estimates using the least squares method. The coefficient obtained is finally inverted to represent the value in degrees per 1 dB (i.e inverse ILD slope). Thus, the result provides the measure of angular resolution in ILDs at various frequencies. Smaller angular resolution implies that the magnitude of ILD increases faster as azimuth increases, and hence the ILD is more sensitive to the azimuth.

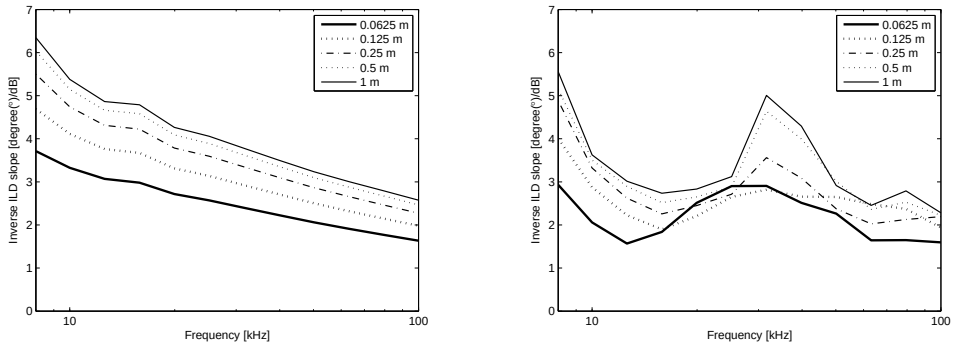
By applying the sphere model analysis, the inverse ILD slope was calculated over the frequencies from 8 kHz to 100 kHz for distances from 0.0625 m to 1 m. Then, the result was compared to that of the measured data. For all 5 distances, higher frequencies show smaller inverse ILD slopes which implies higher sensitivity of ILDs on the azimuth. In other words, smaller angle difference is required to obtain a 1 dB difference of ILD in higher frequencies. This concept is shown in Figure 5.20. Assuming there are two identical sound sources at the same distance but different azimuths, the ILDs change more sensitively over high frequencies whereas the same ILD change is obtained for larger azimuth difference over low frequencies. For the sphere model prediction, the inverse ILD slope decreases as the distance decreases, however the values at lower frequency decrease faster than at higher frequencies.

The measured data are specifically compared to the simulated data as shown in Figure 5.21. The low frequency characteristics in the measured data are shown to be similar to the simulated data. However, the measured value of inverse ILD slope is boosted in the intermediate frequency range between 30 kHz and 40 kHz and this effect is distributed more widely for the distance of 0.0625 m. The frequency region around 30 kHz and 40 kHz is where the spectral notch occurs due to the geometry around the pinna. Also the unusual feature shown in the result measured at 0.0625 m is considered to be due to the measurement limitation in the near field as described earlier. The change in the ILD slope with the distance at the high frequencies in the simulation data is not consistent in the measured data. As the high frequency characteristics are more sensitive to the geometrical feature of the bat-head cast such as the



(a) High frequency region (b) Low frequency region

Figure 5.20: Concept of inverse ILD slope for different frequency. It is defined as the angular difference which is required to obtain a 1 dB difference of ILD in the frontal axis between $-30\,^\circ$ and $30\,^\circ$. It has been found that the inverse ILD slope at high frequencies is relatively small (a) and that over low frequencies is relatively large (b). This result indicates that high frequency is more sensitive in terms of angular resolution when ILDs are used to differentiate the location of the sound at the same distance.



(a) Simulated (b) Measured

Figure 5.21: Inverse ILD slopes for both simulated and measured data. The results are shown for all distances from 0.0625 m to 1 m. There are two phenomena observed in the results from the simulation. One is a decreasing in the inverse ILD slope with increasing frequency and the other is decreasing in the inverse ILD slope with increasing source distances across all frequencies. The measured data agrees to the result from simulated data for the low frequency below approximately 10 kHz.

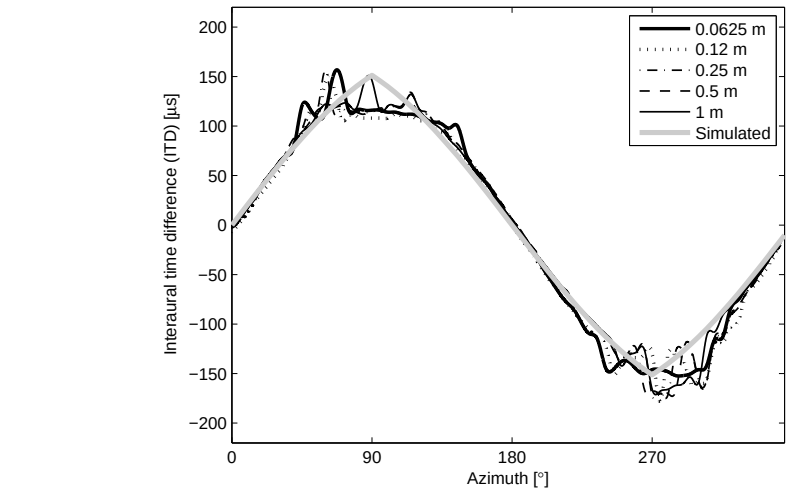
pinna, this confusion is regarded as the effect of the distortion in the high frequency reflection. Hence, the inverse ILD slope measured from the bat-head cast supports the idea that the feature of the head which is different from a complete sphere model, can produce large variability in ILDs at high frequencies where they are naturally sensitive to spectral characteristics due to their short wavelength compared to those at low frequencies.

5.5.5 ITDs

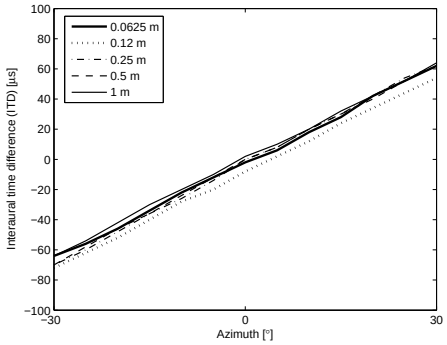
The ITDs are calculated based on the correlation of the envelopes obtained from the HRIRs at both ear. The envelope of the signal at each ear is extracted by using the Hilbert Transform and the delay between peaks of the envelopes at each sound source location is calculated. The details of the ITD calculation are presented in Chapter 4. This process is conducted for the measured data at 5 distances. Also the ITD is calculated by applying the sphere model prediction when the radius of the bat-head is set to 2 cm. In this case, the distance between the source and the head is not included in the calculation as the sound is assumed to be located at infinite distance. The results are shown in Figure 5.22(a). The simulated data appears to have a peak value of 150 μs for maximum ITD when the sound source is located at $\pm 90^\circ$. The measured ITD values appear to be dependent on the azimuth. The absolute value of the ITD increases as the sound source moves towards the lateral positions ($\pm 90^\circ$) and it converges to zero when the sound source is located in the midline. Across 5 different distances, the ITDs are shown to have similar patterns except for the lateral sound location around $\pm 90^\circ$. It is shown that the ITD values varies in an irregular pattern for those lateral sound sources and they are also affected by the change in distance. However, for the sound sources close to the midline, the ITDs change systematically as the angle varies and they also match well to the simulated data. Figure 5.22(b) shows the ITD values at 5 distances for the sources from -30° to 30°. It demonstrates that the ITDs in this region vary approximately from -70 μs to 60 μs in a linear manner for all distances. There is no obvious dependency of ITDs on the distance. According to the result shown in Figure 5.22(c), the ITD change per 1° of azimuth is calculated using 1st order polynomial curve fitting. The Matlab function 'polyfit' was used for this calculation. Hence, the 1st order coefficient of the equation represents the resolution of ITD. Figure 5.22(c) demonstrates the resolution of ITDs for 5 different distances. The average of the resolutions appear to be 2.2 μs and the results from 5 distances are shown to have less than ± 0.1 μs difference from the average value. Also 2.2 μs is estimated to be around 1.5 % of the maximum range of ITDs, which is ± 150 μs.

5.5.6 Summary of binaural findings

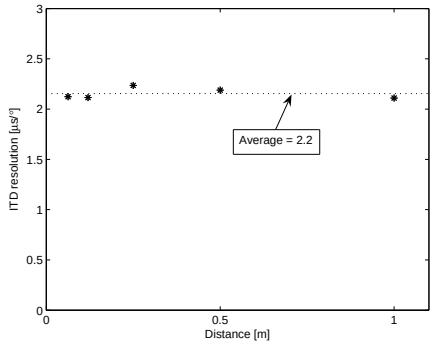
(1) Although the ILDs at the low frequencies appeared to have relatively lower gain compared to those at high frequencies, their acoustical fields in the distal regions (1m, 0.5 m and 0.25 m) change systematically for the lateral sound sources with azimuth (60° - 120°, -120° - -60°) at a fixed distance.



(a)



(b)



(c)

Figure 5.22: ITDs (a) ITDs measured at all distances and at all azimuths (b) ITDs in the frontal axis from $-30\,^\circ$ to $30\,^\circ$ (c) ITD resolution per $1\,^\circ$ at 5 distances

(2) The ILDs at the high frequencies appeared to have relatively larger gain but they do not change significantly as the distance varies in both proximal and distal regions. On the other hand, the ILDs at the low frequencies increase with decreasing source distance and this feature is prominent in the proximal region. The result provides insights into the potential use of different frequency ranges depending on the characteristics of HRTFs according to distance. The low frequency binaural cue is available in distance perception during echolocation and this effect would be more prominent in the proximal region which has substantially larger gain. On the other hand, a high frequency binaural cue is expected to provide robust gain from distal to proximal region, which

enables better detection of the target even in poor SNR environment.

(3) Comparison of results from the sphere model and the bat-head cast shows similar characteristics of HRTF between the measured and simulated data. According to both measured and simulated data, the inverse ILD slope (angular sensitivity) which is calculated for the azimuths between $-30\,^\circ$ and $30\,^\circ$ decreases as the distance decreases and this effect is significant at the low frequencies. At the same time, the simulated data from the sphere model has shown that the inverse ILD slope decreases as the frequency increases, hence high frequency is more sensitive in terms of angular resolution when ILDs are used to differentiate the location of the sound at the same distance. Similar results were shown in the measured data but it was only limited to the low frequency region due to the physical features of the bat-head cast.

(4) The ITDs are relatively independent of distance, however it is expected to be possible to predict angular location of the sound source when the ITD value is given in frontal azimuth for the region of $-30\,^\circ$ and $30\,^\circ$ as it varies in a linear manner for all distances used in the study.

5.6 Discussion

Both monaural and binaural characteristics are analysed in this chapter to investigate the range-dependent characteristics of the HRTFs of the bat-head cast. First of all, as expected based on acoustic theory, the result shows that the acoustic field around the head depends on the size of the head, distance and frequency. The general pattern characteristics measured at each ear were shown to be similar between various distances from 1 m to 0.0625 m. It is therefore expected that the basic characteristics of binaural hearing in bats can be obtained from the distal to the proximal region as long as the magnitude of HRTF is obtained within a satisfactory SNR condition. However, there exists differences between the data measured at different distances in terms of details in the characteristics. Firstly it is observed that, as distance decreases from 1 m to 0.0625 m, relatively more energy at low frequency (20 dB) is gained at each ear than at high frequency (10 dB). This result shows that the gain of HRTF does not increase linearly across all frequency regions from low to high frequency. This implies that one may obtain relatively larger gain in low frequencies when the HRTF is measured at close distance to the sound source. This study also shows that the change in low frequency gain according to varying distance is consistent in the frontal axis, whereas the change in high frequency is only prominent for ipsilateral source positions ($60\,^\circ$, $90\,^\circ$). This

result provides insight into the potential role of low frequency and high frequency in monaural hearing. The monaural characteristics at the low frequency can be used as a cue to perceive the distance for the sound source located in the frontal axis, however, the high frequency cue is only available for ipsilateral sources.

Secondly, the measured data have revealed that the change in distance affects the ILDs in the HRTF and the result is shown to be systematic at low frequencies. The increase in gain in the low frequency regions of 8 kHz, 10 kHz, 13 kHz and 16 kHz appear to be clear when the distance decreases, whereas the gain in the intermediate and high frequencies has shown no obvious change according to the distance. It is also noted that the octave band analysis suppresses the change in the high frequency and it contributes to maintaining the change in the low frequency regions. This study has demonstrated the acoustic ILD field for all distances around the bat-head cast and the results have shown that the ILDs at higher frequencies change faster in general compared to the ILDs at lower frequency when the sound source is located in the front and back. It is also observed that the high frequency generally produces a large gain of ILD for lateral sources and this effect is consistent across the distances. On the other hand, the ILDs at low frequency for lateral sources are shown to be smaller than at high frequency. The source positions, which are particularly close to the interaural axis produced even smaller gains in ILD. A high gain in ILD implies better performance in detection of a target and it perhaps can lead to better characterisation of spectral features. According to the results shown in this study, the binaural processing at high frequencies has better detection performance in lateral sound sources, whereas low frequency might not be beneficial for the detection of the sound both in front/back sources and lateral sources. However, low frequency is considered to have a potential role of contributing to sound localisation of lateral sources because the ILDs at low frequency change systematically not only in front/back sources but also in lateral sources. It is clear that the change of ILDs at high frequency according to the azimuth are relatively small, particularly for lateral sources even though the gain is higher than at low frequency. It can be concluded that the function of binaural signals varies depending on the frequency region and the position of sound source.

Further analysis shows that the ILDs at low frequency change systematically as the distance changes whereas the ILDs at high frequency tend to maintain the magnitude without substantial change according to the distance. Brungart *et al.* have already shown that the low frequency can function as a distance cue for the perception of sound localisation in human HRTFs [Brungart and Rabinowitz, 1999]. Obviously the

'low frequency' in previous study denotes the frequency around 500 Hz to 1 kHz considering the human audible range which is from 20 Hz to 20 kHz. In our case for the study of the bat, the 'low frequency' ranges around 10 kHz. It is notable to find that similar characteristics is shown in the HRTF of the bat. Although the result is predictable as HRTF is purely physical depending on the size of the head, frequency and distance, it provides different aspects of looking at the role of 'low frequency' in bats. The relationship between the distance and the low frequency is even more interesting because echolocation is carried out when bats change distances during their flight.

The results from the analysis of angular sensitivity in ILDs imply that a dynamic change in the frequency range of echolocating calls might potentially have functional implication related with the binaural information in ILDs when a bat changes its position while approaching a target. Figure 5.23 describes the sensitivity of ILDs at low and high frequencies, at fixed distance and at moving distance respectively. Using the ILD cues, it is expected that higher frequency will result in smaller angular resolution whereas lower frequency will perform larger angular resolution using the same variation in ILDs. On the other hand, the observed angle between two objects as shown in Figure 5.23 varies as the head approaches the target. As the distance reduces, the relative angle increases. Applying the relationship between the frequency and the angular resolution described in section 5.5.4, the distance effect on the frequency can be interpreted as follows. It is expected that the performance of high frequency to localise an object according to the ILD cues at far distance, can be equally recreated by the lower frequencies at nearby distances. Here, we define 'sensitivity calibration' as a phenomenon which adjust the angular resolution of ILDs to be constant as a receiver moves towards the target. Therefore, it can be concluded that the calibration of the angle separation is achieved by lowering the frequency when the head approaches the object.

The new finding about the range-dependent HRTF in the bat-head cast in this study is the calibration of angular sensitivity in ILDs. Some bats are known to increase their call bandwidths as they approach their prey and this adaptive change is also found when they approach an obstacle during flight [Kalko and Schnitzler, 1993; Moss et al., 2006]. In general, broadening the bandwidth enables better estimation of target range and an increase in frequency helps the extraction of object features [Schnitzler et al., 2003]. Holderied and von Helversen [2006] have also shown that *Myotis mystacinus*, which uses broadband FM call, adopts range-dependent call design during flight to localise nearby obstacles accurately. Therefore, it has been generally accepted that an adaptive change

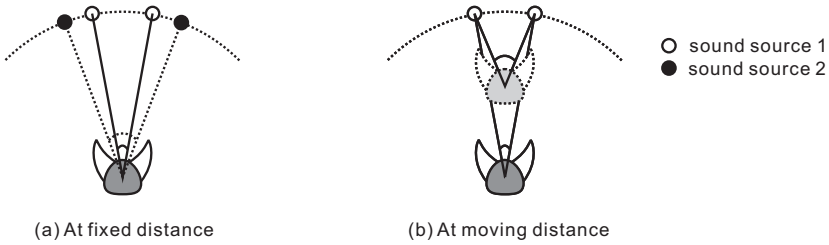


Figure 5.23: ILD differences between two separated sound sources at fixed distance (a), high frequency performs high sensitivity in the separation of two different object, whereas the low frequency characteristics requires relatively larger angle separation to achieve as same amount of magnitude difference as in high frequency. At moving distance (b), the relative angle towards the two different targets increases as the head approaches them.

in call structure enables better performance in echolocation. The results shown in this study support the same idea by providing an acoustical evidence based on the HRTFs of a bat-head cast. This study has found that the binaural processing might contributes to the accurate localisation during flight by adopting an adaptive echolocating call signal structure which is extended towards lower frequency. The results have shown that the rate of changing ILDs in the frontal axis from -30° to 30° varies depending on both frequency and distance. It is expected that the acoustical angular sensitivity of ILDs will vary as the distance between the head and the sound source varies when the frequency is fixed. It has been shown that the sensitivity can be calibrated by lowering the frequency when the head approaches the sound source. Therefore, this result potentially explains why some species of bats lower their frequencies during their echolocation. The calibration of angular sensitivity shown in this study is still required to be examined by behavioural studies in bats, however the acoustical characteristics shown in the HRTFs of a bat head cast shows possible link between the behaviour of lowering the frequency and its functional implication of binaural calibration. It is also important to note that the bat-head cast used in this study is modelled based on *Rogettus aegyptiacus* which uses broadband click signals for echolocation. Holland et al. [2004] have demonstrated the signal characteristics of an impulse click in terms of duration, amplitude and frequency spectrum which might have implications for obstacle avoidance during echolocation in *Rogettus aegyptiacus*. As obstacle avoidance includes a motion of approaching an obstacle and perceiving its location, the binaural calibration effect shown in this study provides a new aspect of signal characteristics which imply a contribution to maintaining the localisation performance during a bat's flight.

Also, although the results imply an advantageous function for low frequency with decreasing distance, it should be noted that the exact distance from a target which a bat starts adopting and utilising the lower frequencies has been rarely discussed in the literature. This distance can be only verified by investigating the flight path and speed along with the recorded signals. Based on the sonograms of echolocating signals shown in previous studies, it is expected that the position of a bat where it lowers the frequency of echolocating calls is near to an object. The results shown in this study suggest the importance of proximal region (less than 20 cm for a head size of 2 cm in radius) where the characteristics of low frequency change effectively. Therefore, it is considered necessary to investigate an adaptive change in emitted signals, particularly in this region in future work.

On the other hand, as shown in the analysis, the angular region from -30° to 30° was chosen to calculate the angular sensitivity. However, one can argue about the effect of changing the angular region and defining the angular resolution depending on the analysis region in the frontal axis. Here, a smaller angular region from -5° to 5° is applied to the analysis in the sphere model and compared to the results from the analysis of the region from -30° to 30°. Figure 5.24 plots the frequency over the ILD variation between the specified angles. First of all, higher frequencies appear to have larger ILD differences for both cases. As expected, larger ILD difference is found in the wider angular region (from -30° to 30°) than in the narrower angular region (from -5° to 5°), however it is interesting to note that the effect of distance is also larger in the wider angular region. This is considered to be due to the fact that the relative angular change between two positions according to the change in distance is larger in wider angular region. Figure 5.25 demonstrates the perceived angle (θ_p) between two sound sources located symmetrically off the centre when the distance changes from 1 m to 0.1 m with 0.01 m of resolution. The initial reference distance was set to 1 m and and the initial azimuth (θ) is defined as the angle with reference to the midline. Eq 5.4 was derived from the geometry of the head and the source positions to calculate the perceived angle as follows.

$$\theta_p = 2tan^{-1}\left(\frac{sin\theta}{cos\theta + d - 1}\right) \tag{5.4}$$

where d denotes the distance between the source and interaural centre. The result shows that the initial separation angle from -30° to 30° causes not only larger perceived angle but also larger change in the angles as the distance reduces compared to the case of -5° to 5°. However, it is clear that the change in frequencies (i.e lowering frequency) is required to achieve same ILD variation between two angles during the

decrease in distance for both cases. Figure 5.24 also illustrates the example of lowering frequency to maintain the constant ILD variation for both angular regions. The ILD variation value appear to be different between two different regions as expected. However, it is noted that the range of frequency which is required to be lowered to maintain the ILD variation is shown to be similar between the two cases.

Lastly, compared to the ILDs characteristics in different ranges, the envelope of ITDs were shown to be relatively independent of distance. It is noted that the ITDs in the frontal axis from $-30\,^\circ$ to $30\,^\circ$ change linearly and the characteristics were almost same for various measured distances. This implies that the ITDs can be used as a cue to locate the sound along the horizontal line when the sound source is close to the midline. This implication is purely based on the acoustical data of the signals received at two ears but the empirical data shows the potential use of ITD as a distance-free cue for sound localisation.

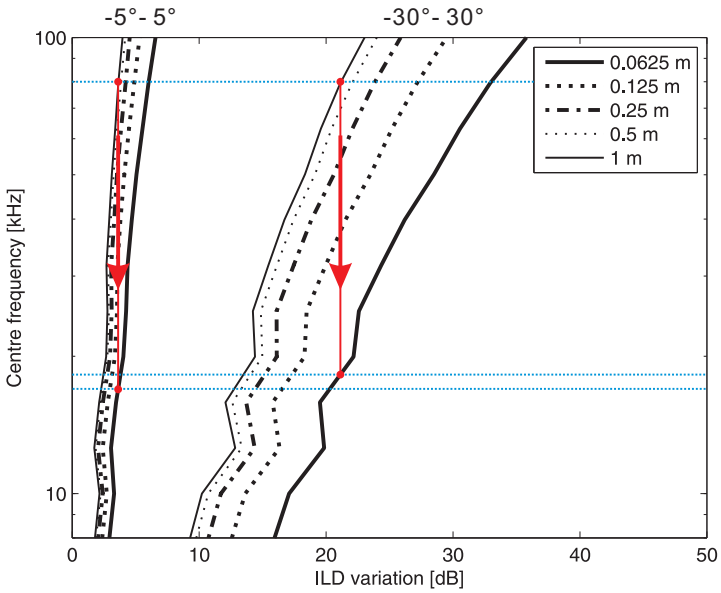


Figure 5.24: The relationship between ILD variation and frequency according to the range of frontal source positions ($-5\,^\circ$ to $5\,^\circ$ and $-30\,^\circ$ to $30\,^\circ$) and distances.

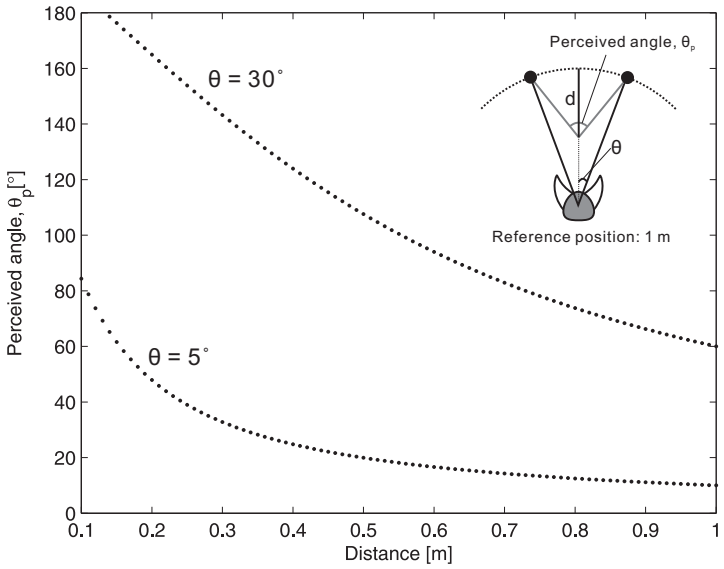


Figure 5.25: Perceived angle (θ_p) vs. distance (d)

Chapter 6

Discussion

6.1 Basic understanding of binaural processing in bio-inspired technology

The results of this research have shown the binaural characteristics found in the bat-head cast to reveal the principles of adaptive change in the echolocating signals when a bat approaches a target. Measured head-related transfer functions (HRTFs) at various distances from the distal to the proximal region provide an insight into the involvement of binaural hearing in echolocation. It is already well established that the interaural difference cues that are used in passive hearing localisation are also of use in determining horizontal azimuth using echolocation. However, it is unknown whether there are other benefits of binaural hearing beyond the utilisation of interaural cues to locate prey and objects. The binaural specialities under the echolocating condition have rarely been revealed. The echolocating condition is associated with active sensing behaviour such as adaptive change in calls, flight, movement of head and pinna which vary during echolocation. This study has adopted a 'distance' factor to investigate its effect on binaural hearing in echolocation.

As the importance of bio-inspired technology is increasingly being accepted, bat echolocation is attracting growing research interest, particularly in designing sonar receivers [Müller and Kuc, 2007; Waters, 2007]. The processing of interaural disparities between the ears has been adopted in applications such as robots that use sound localisation techniques [Walker et al., 1998; Carmena and Hallam, 2004]. However the application of binaural transducers in this technology is still under-utilised as a lack of fundamental understanding of binaural processing has restricted its adoption. Most binaural technology linked to echolocation or sonar has used very basic applications of ILDs and ITDs at fixed positions of sound source and receiver.

In biological research which deals with living bats, the investigation of sensorimotor characteristics in bats is regarded as a challenging task because of the natural difficulties in the measurement of echolocation signals generated by flying bats. Therefore, no empirical measurements of binaural signals reaching the ears of a flying bat have been made yet. Thus, the current study has taken a novel approach by investigating the range-dependent binaural signals in echolocation based upon measurement of the HRTFs of a bat-head cast.

Based on this study, there are three factors which can potentially be applied to the design of binaural processing in a sonar system. They are the presence of the head in the binaural receiver, frequency range and distance. Firstly, it is already well known that HRTF is defined by the head and pinna and it removes the ambiguity of the location of sound sources on the cone of confusion [Blauert, 1997]. It is interesting to note that, whereas the pinna of the bat has inspiring many designers of sonar systems, the effect of the head has seldom been considered. Interest in the pinna may be because its shape and directionality manipulates sensitivity to the spatial intensity of sound. Also, many bats possess unique and unusually shaped pinnae which have inspired receiver design [Carmena and Hallam, 2001; Müller, 2004]. Thus, signal characteristics modified the by head itself have received little attention. Furthermore, a binaural sonar is widely defined as having two microphones and one signal generator without the model of the head [Huang et al., 1997; Walker et al., 1998; Kuc, 1997a, 2002, 2007; Shi and Horiuchi, 2007]. Although there is no head/pinna presence, a binaural sonar which has any geometrical shape may produce a transfer function for a given direction of sound source. Holderied and Helversen have recently measured the binaural echoes from an object using an artificial bat-head consisting of one loudspeaker and two microphones and the results have shown that the binaural transfer function of the object echo provides a potential indicator to discriminate the object orientation [Holderied and von Helversen, 2006]. This result suggests that the evaluation of transfer function of a sonar head might have an importance on improving the performance echolocation. In general, the presence of a head in a shape of sphere not only provides cues to remove the ambiguity of sound location but also functions as a low-pass filter. Therefore, this idea can be expanded to develop a receiver which reduces the artefacts caused by high frequencies.

In addition, knowing the dynamic characteristics of the HRTF provides feedback that is of use in understanding receiver performance with sound sources at different locations. This study also provides insights into the design of an optimal echolocating signal. The frequency characteristics of the received signal are correlated with the characteristics

of the receiver. For example, the returning echo is filtered by the characteristics of receiver and the filtering characteristics are determined by its size and shape.

The HRTF can be broadly divided into low and high frequency characteristics. The effect of low frequencies is relatively straightforward, showing a systematic change according to azimuth. The role of low frequency in echolocation has thus far been underestimated. Thus, the current study provides a new aspect of signal design.

Finally, increasing distance not only attenuates the returning echo but also changes the characteristics of the HRTF. Again, low frequencies are more affected by distance than high frequencies due to the diffraction around the head. As the change in the characteristics of low frequencies according to the distance is systematic, there is potential to use low frequency as a distance cue [Brungart and Rabinowitz, 1999]. Furthermore, the binaural processing of low frequencies effectively contributes to maintaining the sensitivity of ILD in the frontal axis when the receiver moves towards the sound source. Therefore, this principle is expected to provide calibration cues in sound localisation by using low-pass filtered signals at varying distances.

6.2 Practical issues in designing binaural receivers

Designing binaural receivers may mean more than using two sensors. Waters [2007] has insisted that the important problem in the bio-inspired sonar system is a production of repeatable, accurate and high power acoustic signal, which is similar to bats. It can be said that it is also important to build a robust and reliable receiver which can perform as bat's ears. Selecting suitable microphones and modelling the receiver may require accurate validation before it is used in a system application. A preliminary step is to note the acoustical properties of the measurement system. The available frequency range should be determined based on the SNR properties of the system. With binaural receivers, it is particularly important to ensure that the SNR covers a larger range of ILDs than is expected. Such an ILD range can be roughly predicted using the sphere model calculation if the size of the head and the frequency is provided [Duda and Martens, 1998]. The system properties are based purely on the signal generator and the sensors. It is recommended that the frequency characteristics are validated when the transmitter and receiver are connected. Microphone directionality can also affect the designing of the receiver. In this study, a 1/8-inch pressure field microphone is used. Due to its small size, it is considered suitable for the purpose of taking random-incidence measurements when the frequency response is less dependent

on angle of incidence. However, if the loudspeaker has a strong directionality pattern, it should be noted when the measured data are analysed as the gain of received signal might vary depending on the angular direction of the loudspeaker.

It is vital that the central position of the interaural axis is set up accurately. As the binaural receiver function relates to the geometry and the distance between two different receivers, it is easy to produce measurement error when the central position of the interaural axis is not accurately adjusted. In general, there are two methods to adjust the interaural axis. One is based on the geometry of the rig and the other on measurement. Geometrical adjustment uses visual images such as laser projection, photography or simple experimental observation. It may provide a visually correct result but there may be also error which is introduced by a false reference position. This study has adopted the visual references on the measurement rig which consists of metal beams. The metal beams have been useful to position the source generator and the receiver, and move them systematically. Hence, the visual referencing point was able to be fixed. On the other hand, the measurement method uses the measured signals received at two microphones and the experimenter adjusts the position of source generator and receiver until the measured data correspond to the predicted result. For example, the ITDs for the sound source at $0\,^{\circ}$ is expected to be zero as the arrival time of the sound is the same at each ear. Hence, a particular position of the receiver gives feedback to the analysis stage and the position is corrected until the analysis result reaches the correct value. In this study, the updated rig with fixed structure described in Chapter 5 gave better accuracy than the earlier rig described in Chapter 4. It has been found that small errors between the signal generator and the receivers can produce substantial measurement error particularly in the high frequency responses which are relatively more sensitive. Based on this study, it is recommended to build a rig which the position of the source generator and the receiver can be systematically adjusted and then to examine the error using the measurement method.

Measurement of the HRTF of echolocating bats is an accepted method for revealing the binaural cues used in their echolocation, however this study has investigated the effect of measurement on the measured data under the condition which a bat might experience during its echolocation such as a change in distance. As mentioned in Chapter 4, there is no standardised distance designed for such measurement. The result of this study demonstrates that there may be significant differences in measured HRTF functions as the distance between the sound source and the receiver is altered. The study predicts that low frequencies would be affected relatively more than high fre-

quencies as the magnitude of the gain would be higher in this range. High frequency characteristics are also expected to vary especially for contralateral sound sources as distance varies. It can be concluded that consideration of the distance parameter is of great importance when a binaural receivers are characterised by measuring its HRTF.

6.3 Auditory processing and binaural receivers

The auditory model inspired by bats and its analysis of target echoes has the potential to recreate the psycho-acoustical environment that a bat might experience during echolocation. Auditory analysis is useful to represent both spectral and temporal characteristics. For example, it has been shown that the spectrogram correlation and transformation (SCAT) model is advantageous for reconstructing the temporal acoustic image [Simmons et al., 1995; Saillant et al., 1993; Peremans and Hallam, 1998]. In particular, analysis of the temporal characteristics of returning echo is considered necessary to separate two glints from a target and it is regarded as an important task in echolocation to detect the target range. It might be necessary to extract a temporal auditory image rather than the frequency characteristics due to the fact that the series of echoes or the relationship between the emitted signal and the target echo can be only represented in the time-domain. On the other hand, most of the auditory processing models in bats have been adopted and designed based on standard gammatone filter banks. The advantages of gammatone filterbanks are that they can be modelled relatively easily, based on the centre frequency and the bandwidth [Lyon, 1996], and that they resemble the temporal filter characteristics of the basilar membrane in mammalian cochlea. Because of their convenience, gammatone filterbanks have been used to model the auditory processing in several bats such as *Rhynolophus ferrumequinum* and *Phyllostomus discolor* [Müller and Schnitzler, 2000; Wiegrebe, 2008].

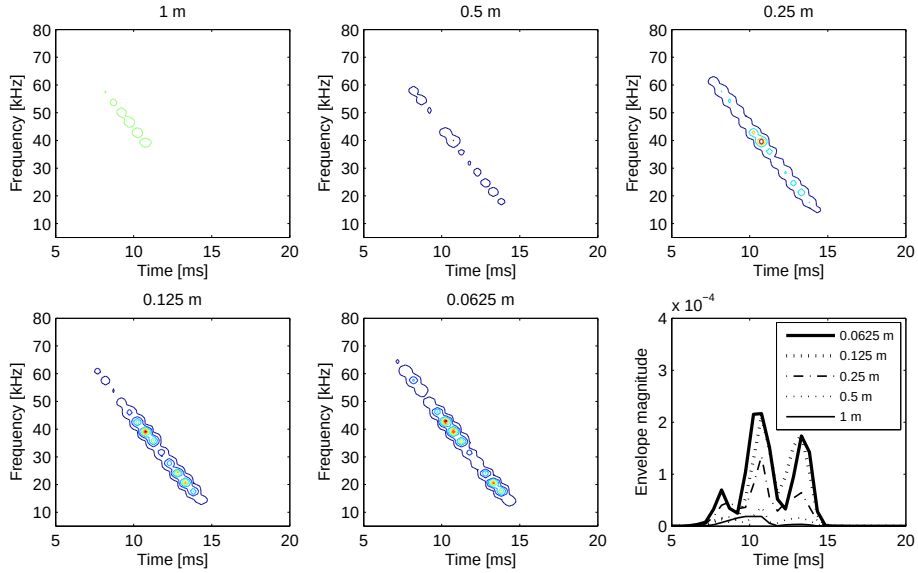
The model in this study gives much flexibility in designing bat-inspired auditory processing techniques. By changing Q factors and using different types of signals (CF/FM), the same target impulse response can produce significantly different auditory spectrograms. A higher $EarQ$ value is advantageous to avoid leakage of the energy contained in each frequency band and overlapping of gammatone filterbanks between adjacent centre frequencies. When the envelope of the auditory spectrogram is extracted, the shape is similar to the envelope of the original impulse response in the time-domain. However, for example, irregular fluctuation in the impulse response's envelope can be reduced in the auditory spectrogram by increasing the $EarQ$ factor. The result helps to enhance the peak characteristics of the envelope in particular when the FM signal

is used for echolocation. Müller and Schnitzler have suggested that modelling of the auditory system of CF-bats provides an opportunity to separate the strongest echoes from a single target [Müller and Schnitzler, 2000], and hence it reduces distortion in the demodulation. In this study, a similar effect was shown when CF and FM signals were simulated for object discrimination. However, it is interesting to note that the thickness difference was suppressed and that size information became more dominant when auditory processing was applied to extract the envelope magnitude in the echo spectrogram. On the other hand, the thickness characteristics were found to be contained in the auditory spectrogram when the FM signal was used for echolocation. Note that the auditory spectrogram of the CF signal does not produce a clear thickness difference as shown with the FM signal. This can be attributed to a limitation in the ability to represent spectral characteristics using a single carrier frequency.

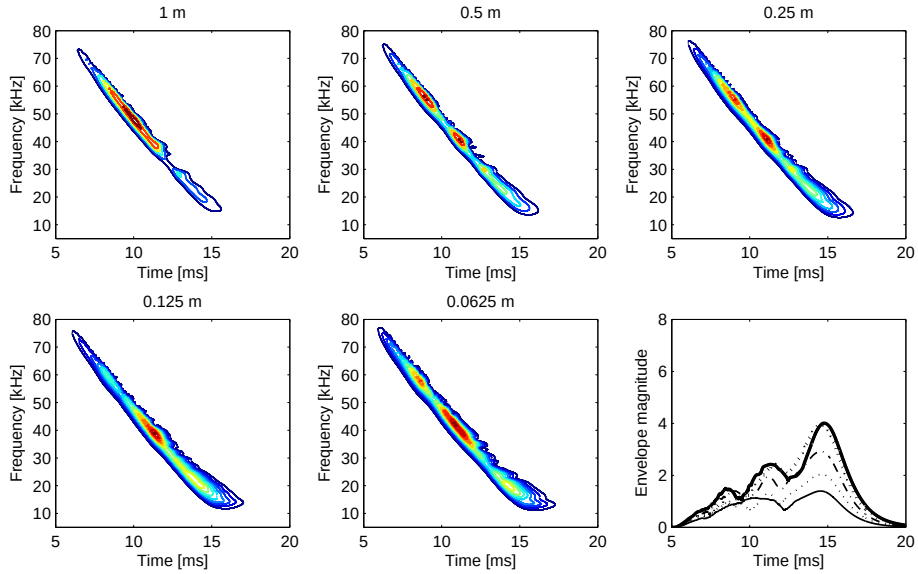
This analysis implies that there is an opportunity to adopt binaural auditory processing as a method to enhance the representation of received echoes to show their acoustic flow. An illustration is provided by comparing conventional FFT processing and auditory processing. To demonstrate the spectral component of signals varying with time, the conventional FFT spectrogram is used. Figure 6.1 shows the difference between conventional FFT processing and auditory processing on the measured HRIRs at five distances from 0.0625 m to 1 m (see Chapter 5). To simulate the echo, each HRIR measured at 30 ° is convolved with the FM downward sweep from 80 kHz to 8 kHz. For the FFT spectrogram, 512-point FFT, Hamming window of 512 samples in length and 50 % of overlapping between time segments are applied. A larger FFT size produces higher frequency resolution but lower resolution in the time domain, while a smaller FFT size has the opposite effect. Because of this trade-off issue, an adjustment in their sizes is required to obtain an appropriate resolution for both frequency and time.

The same HRIRs are processed through the auditory model as suggested in Chapter 5. The number of gammatone filterbanks between 8 kHz and 80 kHz was set to 100 and the $EarQ$ factor was set to 10. The conventional FFT spectrogram and auditory spectrogram are both plotted at five distances and the envelopes are calculated by summing the amplitudes extracted from the frequency bins in each time segment. There are two major areas for comparison between the two processing methods. Firstly, the gammatone filterbanks provide sufficient frequency resolution and preserve the better resolution in time domain compared to the conventional spectrogram. It is observed that the peak and notch characteristics curve is smoother when the gammatone filterbanks are applied compared to the FFT results. Secondly, the auditory envelope

represents the better acoustical characteristics shown in the HRTF at various distances which the low frequency components increase systematically as the distance decreases. The result emphasises the potential use of biologically-inspired signal processing in sonar systems. Improvements such as noise suppression, gain control and flexibility in signal design are expected when the auditory processing is applied. Furthermore, it is worth emphasising that the binaural signals represented by the HRIRs and the output of the auditory processing model, are essential components for simulating further brain processing in order to investigate neurophysiological activity during echolocation.



(a) FFT processing



(b) Auditory processing

Figure 6.1: Comparison between two different processing methods (a) FFT (b) auditory processing

6.4 Spherical model and prediction of the calibration process

The HRTF prediction using the sphere model calculation is a useful method to investigate the effect introduced by the size of the sphere, frequency and distance between the sound source and the receiver. It also helps to understand the data measured from real/artificial heads and the head with various types of pinnae. The result of investigations using models with simple geometry can explain the major features found in the HRTF data from the artificial bat-head, hence it may be useful in designing biomimetic sonar systems. More importantly, the sphere model can be used to predict the performance of binaural hearing in bats which might have implication for the echolocating process. The variation in the magnitude of ILDs for the change of frontal source positions were found to depend on both frequency and distance. This adaptive change is shown to be systematic and significant in the low frequency range. Hence there is a potential to use adaptive ILD cues to calibrate the angular rate of ILD change, which is denoted as the inverse of ILD slope, or angular resolution. This potentially explains the adaptive change in the bandwidth of echolocating signals towards lower frequencies when a bat approaches a target.

The adaptive ILD change can be proved by simulating the sphere model prediction as shown in Figure 6.2. It has been simulated for distances from 0.0625 m to 10 m which covers both distal and proximal ranges, and for a sphere radius of 1 cm, 2 cm, 3 cm, 4 cm, 5 cm and 10 cm. Note that a 2 cm radius corresponds to the size applied in the prediction model used to compare the measured data from the artificial bat-head cast in Chapter 4 and Chapter 5, and a 10 cm radius corresponds to the approximate size of a human head. Three points can be noted. Firstly, the inverse ILD slope is generally smaller for a larger radius of the sphere which implies that a larger sphere will achieve a larger range of ILDs when the frequency range is fixed. This is due to the fact that the acoustical wave travels a longer distance to reach the distal ear from the sound source when the sphere is larger. Hence acoustic gain is reduced. Secondly, the results show that the inverse ILD slope decreases as frequency increases but it does so more slowly as the radius increases which means that smaller spheres introduce greater dependence on the frequency. Thirdly, the inverse ILD slope decreases as the distance decreases and the effect is larger in lower frequencies, except the case when the sphere radius is 1 cm. In that case, the frequency range between 10 kHz and 20 kHz has shown a larger inverse ILD slope which implies smaller difference of ILDs in the defined angles. This is considered to be due to the diffraction of those frequencies around the sphere which

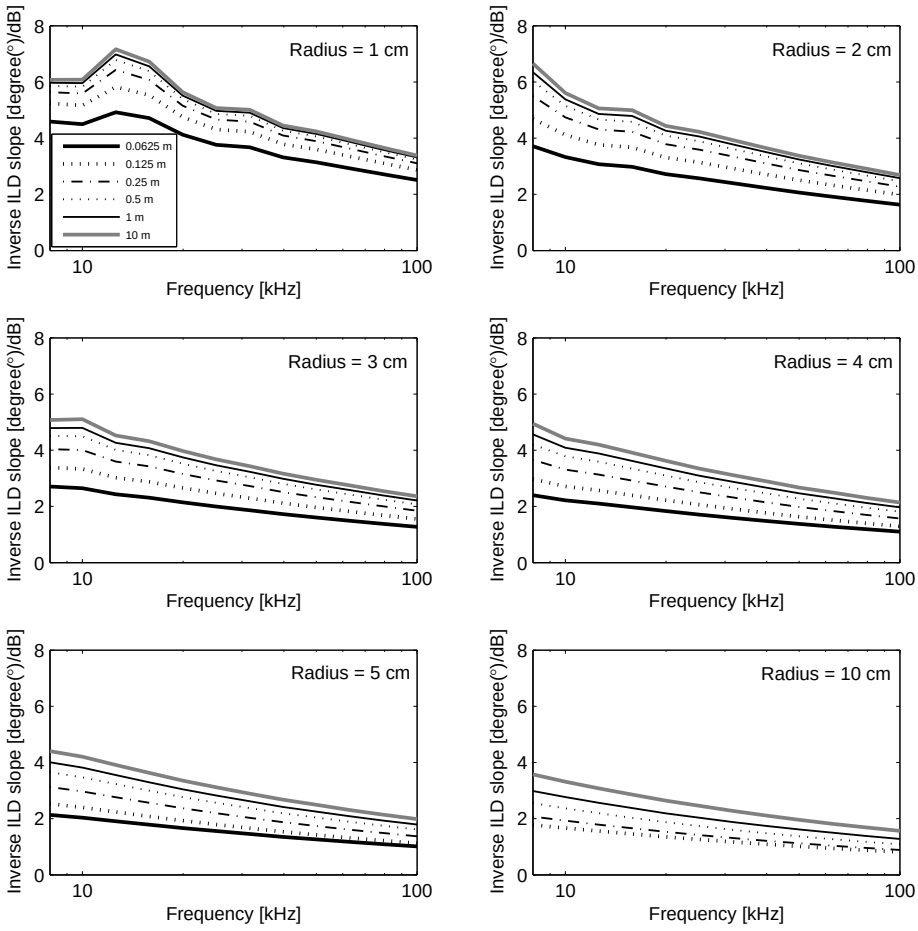


Figure 6.2: The simulation of sphere model prediction for various sizes of the head. The result is presented for distances from 10 m to 0.0625 m between the head and the sound source.

reduces ILDs. The overall results predict the performance of the binaural receiver when it is designed in the shape of a sphere with rigid materials. It is interesting to note that a larger sphere, for example, one with 10 cm radius, is expected to achieve better angular resolution over distance compared with smaller spheres. In general, the size of the sphere determines the range of possible ILDs when the frequency range is fixed and the binaural performance varies as the distance changes. On the other hand, as shown in Chapter 5, it is expected that non-spherical characteristics such as the shape of frontal face, size and shape of the pinna will alter the results of the prediction. The bat-head cast in this study resulted in an increase of the inverse ILD slope in the intermediate and high frequency regions (Chapter 5).

Here, based on the sphere model prediction result, the frequency range which is required to maintain 3° of inverse ILD slope during the change in distance from 10 m to 0.0625 m is calculated. 3° of angular calibration was chosen by the author to enable a comparison between different sizes of spheres provided in the simulation. Note that the horizontal angular resolution of localisation in bats has been found to be around 1.5° as reviewed in Chapter 2. Figure 6.3 demonstrates the frequency component which produces a 3° inverse ILD slope at each distance. For the values of inverse ILD slopes at all frequencies and distances, those values which range from 2.85° to 3.15° are detected and plotted, hence 0.3° of error is tolerated in this calculation. The defined tolerance level is considered to be acceptable because it is smaller than the minimum discrimination angle found in bats, as mentioned earlier. As a result, it clearly shows that a certain bandwidth of the frequency range is required to maintain the calibration process as distance varies. It should also be noted that calibration is only achieved by lowering frequency when distance decreases. In addition, it was found that a lower frequency range was required for a larger sphere compared to a smaller one when the angular resolution of calibration was fixed. Finally, Figure 6.4 plots the maximum and minimum frequencies found in Figure 6.3. As the radius of the sphere increases, the maximum frequency decreases from 100 kHz to less than 20 kHz and the minimum frequency decreases from 50 kHz to less than 10 kHz. Hence, the bandwidth of the frequency range changes from approximately 50 kHz to 10 kHz depending on the size of the sphere. The result reveals that there exists optimum size of the binaural receiver and the frequency range to utilise the binaural performance represented as interaural level differences. At a fixed frequency range, the smaller spherical receiver is expected to require a broader bandwidth of frequency.

As a result of the above analysis, an optimal size of the head for sonar processing can be proposed. Figure 6.5 illustrates the relationship between head size and frequency and potential design options for the sonar head. In the context of evolution in the mammals, it has been already found that a mammal with a smaller head size tends to possess the ability to hear a higher frequency range [Koay et al., 1998]. It is considered necessary for small-headed mammals to use high frequency signals with shorter wavelength to produce interaural level differences to enable sound localisation. Therefore, it can be said that the use of ultrasound up to 200 kHz by a bat is not an extraordinary phenomenon beyond the general features found in the other mammals. In other words, it is regarded as 'natural' to design a small-headed receiver for high frequency and a large-headed receiver for low frequency ranges to utilise the binaural processing. However, if a small-headed receiver were to have a low frequency range, this could be

expected to cause much diffraction around the head, rendering ILDs unavailable. On the other hand, as shown earlier, a large-headed receiver with a high frequency range is expected to be more useful in producing higher sensitivity in ILDs than a small-headed receiver when same frequency range is applied. Therefore, the optimum size of sonar head can be determined based on the level of sensitivity in ILDs that one aims to achieve.

Finally, although the theoretical model presents a prediction of binaural performance, it is important to compare the result with real experiments and applications. It should also be noted that the spherical model represents a pinna-less head. Therefore, it should be expected that there is further potential for altering binaural performance, depending on the pinna attachment and its features.

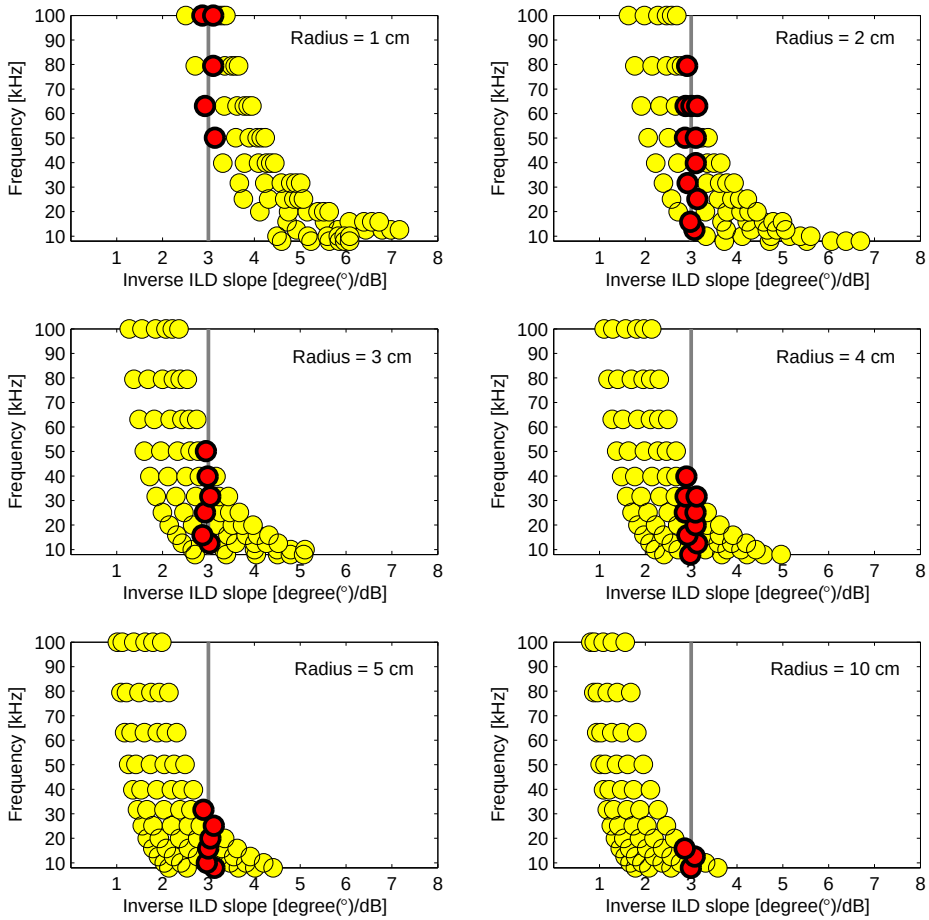


Figure 6.3: The simulation of sphere model for 3 ° calibration effect. The results are shown for all distances from 10 m to 0.0625 m and the frequency regions which have 3 ° of inverse ILD slopes, are detected and shown as red circles in the plots

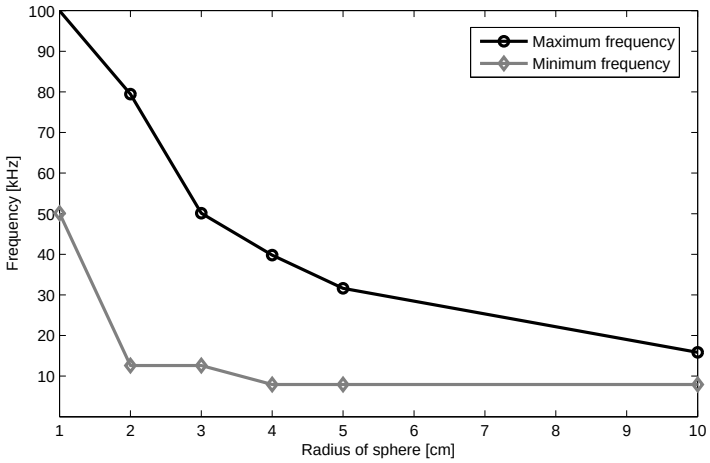


Figure 6.4: Desired maximum and minimum frequency prediction for the ILD angular calibration

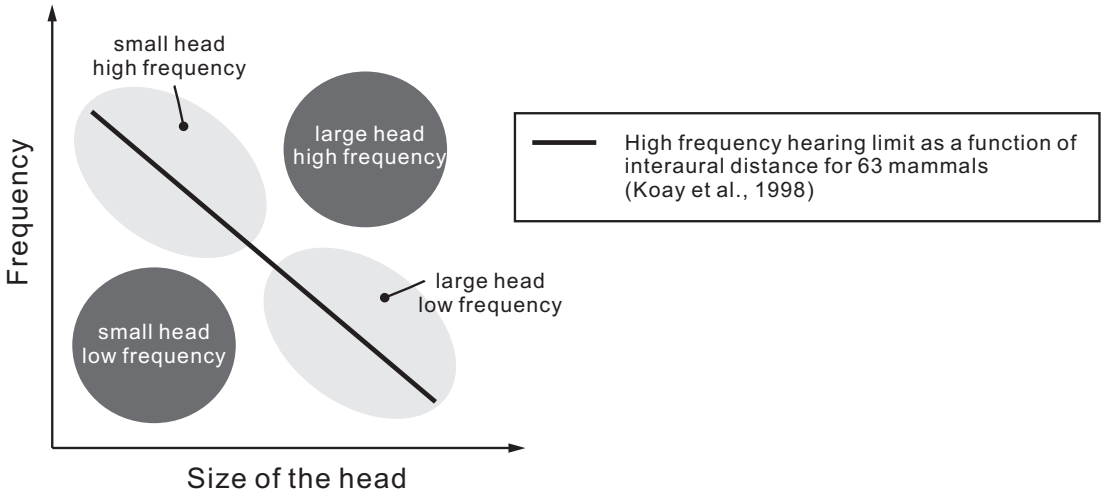


Figure 6.5: The relationship between the size of the head and the frequency

Chapter 7

Conclusions and future work

7.1 Conclusions

First, the present study investigated object discrimination in the echolocating process by using the mathematical convolution model. The model was shown to be advantageous as each echolocating stage such as emission, reflection and reception can be individually simulated under the assumption that each component is a linear time-invariant system. The object discrimination was successfully applied to the model by simulating the echo from different types of discs. The spectrogram analysis from the results of the various targets showed the different characteristics depending on the types of echolocating signals. The magnitude in the spectrogram using the CF signals changed depending on the size of the target but no spectral features such as peaks and notches were observed. However, the FM signals were shown to carry spectral peaks and notches, and the spectral characteristics were more prominent as the sweeping ratio of the signal increased. In particular, the concept of the 'activated frequency channel' was adopted when FM signals were used. Hence, the number of frequency bins exceeding the threshold of the noise level was calculated for each target. The results showed a high correlation between the mean number of activated frequency channels and the diameter of the object. Also the variance of the activated frequency channels across the time domain showed a high correlation with the sweeping ratio of the signal. Therefore, it can be concluded that this analysis tool provides an alternative parameter in terms of the size discrimination using FM signals. Further study has been carried out to extend the echolocating model to include the hearing system of the bats. The general mammalian auditory model was applied when the FM signals were used. Also, the specialisation in the cochlea known as the acoustic fovea observed in CF bats was taken into account for the modelling when a CF signal was used. The implication of the results from the modelling supported the view that CF signals are advantageous for Doppler shift compensation whereas the FM signals are advantageous for object

characterisation.

Further investigation was carried out to extend the study to the binaural model using two receivers. This study investigated practical methods for measurement of the HRTF of a bat-head cast and finalised the robust set-up for the hardware environment. The microphone set-up is critical to measure the HRTF without the amplitude distortion in frequency response. The custom-made extension tube was examined and it appeared to be less reliable due to the resonance in the tube. It also limited the upper frequency range of the system. Therefore the right-angled microphone, which has the stable frequency response ranging from 5 kHz to 100 kHz, was finally chosen for the measurements. The measured HRTFs have shown that the peak and notch characteristics in the high frequency response above 10 kHz are vulnerable in terms of repeatability when there is no direct path between the sound source and the receiver. It is noted that the microphone used in this study was assumed to be independent of the angle of sound incidence.

The major results from the binaural model lead to conclusions about the distance effect on binaural echolocation. The gain of the measured HRTFs increases as the source distance decreases and the monaural gain is higher for the lower frequencies by approximately 10 dB. The variation of ILDs according to the change of distance is significant and systematic across the frontal azimuth at low frequencies below 10 kHz, hence the results confirm that the low frequency monaural gain of the HRTF can be used as the distance cue during echolocation. In addition, the ILDs increase as the distance decreases and the effect is prominent at proximal region where the source distance is shorter than 0.125 m. Also the pattern change is smoother at low frequencies than at high frequencies. The effect of an acoustic bright spot was clearly shown at low frequencies, and hence double peaks of the magnitude were shown around lateral azimuths. This result implies a potential role for low frequency in discriminating lateral sound sources.

This study has also shown that the ILD varies linearly for sound sources located at azimuths ranging from $-30\,^\circ$ to $30\,^\circ$. In this range, the ILD variation is more sensitive at higher frequencies when the distance is fixed. At the same time, the ILD sensitivity increases as the distance decreases when the frequency is fixed. Hence, maintaining the ILD sensitivity can be achieved by lowering the frequency of the emitted signal while the binaural receiver approaches the sound source. This result is notable because it may explain why some bats lower their frequencies while they are approaching targets.

For the ILD sensitivity analysis, the measured data corresponded reasonably well with the sphere model prediction at low frequencies although high frequency responses in the measured data were highly affected by the geometry of bat-head cast and peak/notch characteristics induced by the pinna. However, the measured data confirms the role of low frequencies in calibration of the ILD sensitivity. The sphere model prediction shows that the frequency range required to calibrate the sensitivity depends on the size of the head. On the other hand, the ITD were shown to be roughly independent of distance, although the measured value when the source is located at lateral positions is neither consistent nor systematically different across different distances. However, the azimuth is predictable based on the ITD values within the azimuthal range of $-30\,^\circ$ to $30\,^\circ$ which is close to the midline in front.

7.2 Future work

This study motivates further research on the development of an echolocating receiver. As shown in this study, the binaural information manipulated by the existence of the head can potentially contribute to development of more sophisticated sonar receivers. For example, adaptive changes in ILDs according to the movement of the sonar receiver could be applied to improve its binaural performance. Utilising both mobility and binaural information can serve to develop sensory system of aids for the blind or for non-destructive testing of manufactured parts, which clearly requires motion of the reception part of the system. However, the performance of moving sonar needs to be tested under controlled conditions where the experimenter is aware of the acoustical characteristics of the receiver, which can be defined by its HRTF. Then, evaluation of the sonar system is required to be accompanied by the an investigation of the relationship between its performance and the change of its spatial location and frequency characteristics of the outgoing echolocating signal.

On the other hand, the calibrating sensitivity of ILDs, which has been suggested in this study, needs to be examined through behavioural experiments with bats. The result of this study has confirmed the potential use of binaural cues to calibrate the location of a sound source based on the measured data from the artificial bat-head and the computer generated broadband sound source. Binaural recording of echolocating signals received at a bat's ear during flight is required in order to be able to validate the results shown in this study. There have been some studies which have monitored both flight paths and echolocating signals [Holderied et al., 2008]. These studies have been recording the sound by using a fixed microphone mounted inside the experimental

site. The monitoring technique enables systematic changes in echolocating signals to be tracked in relation to proximity of the target. Also there has been a study which has measured a bat's echolocating signal during flight using a telemetry microphone mounted on the bat [Hiryu et al., 2007]. Using these techniques, further measurement of binaural signals during flight is necessary in order to investigate the relationship between binaural hearing and performance of echolocation.

Alternatively, psychoacoustic experiments on human echolocation have potential to help understanding of adaptive binaural information at the perception level. In previous study, a simulation method to virtualise a sound of bat's echolocation and a listening test on human subjects have been proposed to investigate the function of human auditory system on echolocation [Waters and Abulula, 2001]. It is already known that the auditory system of bats resembles that of humans, however how the auditory signals are interpreted in brain processing during echolocation has not yet been revealed. This study has shown in the experimental data that a bat's HRTF is similar to a human's HRTF with frequency scaling effects due to head-size differences. Although there exists differences between bat and human such as hearing threshold, audible frequency range, dimension of the head and types of pinna, human experiments are expected to provide useful feedback about how the mammalian ear perceives binaural information during echolocation.

One of the restrictions in this study is that the HRTF measurement was limited to a horizontal azimuth at 0° of elevation. The future work may expand the study to include vertical localisation. This study eventually aims to reconstruct the HRTF in 3D space. The 3D space can be defined as a function of azimuth (θ), elevation (ϕ) and distance (d), and the HRTF of the receiver at a 3D position will vary depending on those parameters. Once the HRTF in 3D space is reconstructed, the binaural hearing at a given position during echolocation can be predicted based on the spatial location in the path. It is clear that HRTFs of a sonar head in 3D space are considered to be a necessary part in order to understand its binaural performance.

Lastly, the future work also should also aim to develop an optimised signal and auditory processing tool which can provide better performance in echo detection, discrimination and localisation. It is well known that the signal design changes during a bat's search, approach and terminal stage. Previous studies have already confirmed that this signal design changes in a predictable manner depending upon the echolocating task required in each stage. Therefore, designing a signal which changes adaptively in a bio-inspired

sonar system and enabling its practical application remain as challenging tasks. It is also noted that signal processing techniques such as the auditory model, affect the representation of acoustical images of the echo signals. The performance of the sonar task such as gain control and spectral representation can be partly manipulated by those techniques. Therefore, further work is required to develop an effective signal type and its auditory processing which can maximise the performance of practical sonar systems.

References

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. *Journal of the Acoustical Society of America*, 116(6):3594–3605.

Barshan, B. and Kuc, R. (1992). A bat-like sonar system for obstacle localization. *Ieee Transactions on Systems Man and Cybernetics*, 22(4):636–646.

Békésy, G. (1960/1989). *Experiment in Hearing*. New York: McGraw-Hill. (Republished by the Acoustical Society of America).

Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization, Revised edition*. The MIT Press, Cambridge, Massachusetts, London, England.

Boonman, A. and Jones, G. (2002). Intensity control during target approach in echolocating bats; stereotypical sensori-motor behaviour in daubenton's bats, myotis daubentonii. *Journal of Experimental Biology*, 205(18):2865–2874.

Boonman, A. and Ostwald, J. (2007). A modeling approach to explain pulse design in bats. *Biological Cybernetics*, 97(2):159–172.

Brungart, D. S. and Rabinowitz, W. M. (1999). Auditory localization of nearby sources. head-related transfer functions. *Journal of the Acoustical Society of America*, 106(3):1465–1479.

Brungart, D. S., Rabinowitz, W. M., and Durlach, N. (1996). Auditory localization of a nearby point source. *Journal of the Acoustical Society of America*, 100(4):2593–2593.

Carmena, J. M. and Hallam, J. C. T. (2001). Estimating doppler shift using bat-inspired cochlear filter bank models: A comparison of methods for echoes from single and multiple reflectors. *Adaptive Behavior*, 9(3-4):241–261.

Carmena, J. M. and Hallam, J. C. T. (2004). The use of doppler in sonar-based mobile robot navigation: inspirations from biology. *Information Sciences*, 161(1-2):71–94.

Carmena, J. M., Kampchen, N., Kim, D., and Hallam, J. C. T. (2001). Artificial ears for a biomimetic sonarhead: From multiple reflectors to surfaces. *Artificial Life*, 7(2):147–169.

Duda, R. O. and Martens, W. L. (1998). Range dependence of the response of a spherical head model. *Journal of the Acoustical Society of America*, 104(5):3048–3058.

Erulkar, S. D. (1972). Comparative aspects of spatial localization of sound. *Physiological Reviews*, 52(1):237.

Firzlaff, U. and Schuller, G. (2003). Spectral directionality of the external ear of the lesser spear-noise bat, phyllostomus discolor. *Hearing Research*, 181(1-2):27–39.

Firzlaff, U. and Schuller, G. (2004). Directionality of hearing in two cf/fm bats, pteronotus parnellii and rhinolophus rouxi. *Hearing Research*, 197(1-2):74–86.

Forbes, P. (2006). *The Geckos Foot*. Harper Perennial.

Fuzessery, Z. M. (1996). Monaural and binaural spectral cues created by the external ears of the pallid bat. *Hearing Research*, 95(1-2):1–17.

Gardner, B. and Martin, K. (1994). Hrtf measurements of a kemar dummy-head microphone. *MIT Media Lab Perceptual Computing - Technical Report ♯280*.

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(6):1704–1710.

Glasberg, B. R. and Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research*, 47(1-2):103–138.

Griffin, D. R. (1958). *Listening in the Dark*. Yale Univ. Press, New Haven, Conn.

Griffin, D. R., Dunning, D. C., Cahlander, D. A., and Webster, F. A. (1962). Correlated orientation sounds and ear movements of horseshoe bats. *Nature*, 196:1185–1186.

Griffin, D. R., Webster, F. A., and Michael, C. (1960). The echolocation of flying insects by bats. *Animal Behaviour*, 8:141–154.

Grinnell, A. D. (1995). *Hearing in bats: An overview, in A. N. Popper and R. R.R. Fay, eds*. Springer-Verlag.

Grossetete, A. and Moss, C. F. (1998). Target flutter rate discrimination by bats using frequency-modulated sonar sounds: Behavior and signal processing models. *Journal of the Acoustical Society of America*, 103(4):2167–2176.

Grothe, B. (2000). The evolution of temporal processing in the medial superior olive, an auditory brainstem structure. *Progress in Neurobiology*, 61(6):581–610.

Harnischfeger, G., Neuweiler, G., and Schlegel, P. (1985). Interaural time and intensity coding in superior olivary complex and inferior colliculus of the echolocating bat molossus-ater. *Journal of Neurophysiology*, 53(1):89–109.

Hartley, D. J. (1992). Stabilization of perceived echo amplitudes in echolocating bats .2. the acoustic behavior of the big brown bat, eptesicus-fuscus, when tracking moving prey. *Journal of the Acoustical Society of America*, 91(2):1133–1149.

Heffner, R. S. and Heffner, H. E. (1990). Vestigial hearing in a fossorial mammal, the pocket gopher (geomys-bursarius). *Hearing Research*, 46(3):239–252.

Hill, J. E. and Smith, J. D. (1984). *Bats–A natural history*. University of Texas Press, Austin, Texas.

Hiryu, S., Hagino, T., Riquimaroux, H., and Watanabe, Y. (2007). Echo - intesntiy compensation in echolocating bats (*Pipistrellus abramus*) during flight measured by a telemetry microphone. *Journal of the Acoustical Society of America*, 121(3):1749–1757.

Holderied, M. W., Baker, C. J., Vespe, M., and Jones, G. (2008). Understanding signal design during the pursuit of aerial insects by echolocating bats:tools and applications. *Annual Meeting of the Society for Integrative and Comparative Biology*, pages 1–11.

Holderied, M. W. and von Helversen, O. (2006). 'binaural echo disparity' as a potential indicator of object orientation and cue for object recognition in echolocating nectar-feeding bats. *Journal of Experimental Biology*, 209(17):3457–3468.

Holland, R. A. and Waters, D. A. (2005). Echolocation signals and pinnae movement in the fruitbat rousettus aegyptiacus. *Acta Chiropterologica*, 7(1):83–90.

Holland, R. A., Waters, D. A., and Rayner, J. M. V. (2004). Echolocation signal structure in the megachiropteran bat rousettus aegyptiacus geoffroy 1810. *Journal of Experimental Biology*, 207(25):4361–4369.

Huang, J., Supaopngprapa, T., Terakura, I., Ohnishi, N., and Sugie, N. (1997). Mobile robot and sound localization. *Intelligent Robots and Systems, Proceedings of the IEEE/RSJ International Conference*, 2:683–689.

Jen, P. H. S. and Chen, D. (1988). Directionality of sound pressure transformation at the pinna of echolocating bats. *Hearing Research*, 34(2):101–118.

Jones, G. and Holderied, M. W. (2007). Bat echolocation calls: adaptation and convergent evolution. *Proceedings of the Royal Society B-Biological Sciences*, 274(1612):905–912.

Kalko, E. K. V. and Schnitzler, H. U. (1993). Plasticity in echolocation signals of european pipistrelle bats in search flight - implications for habitat use and prey detection. *Behavioral Ecology and Sociobiology*, 33(6):415–428.

Kick, S. A. and Simmons, J. A. (1984). Automatic gain-control in the bats sonar receiver and the neuroethology of echolocation. *Journal of Neuroscience*, 4(11):2725–2737.

Kim, S. Y., Papadopoulos, T., Allen, R., and Rowan, D. (2008). Using a model of bat echolocation to differentiate object dimensions and materials. *Proceedings of Biological Approaches for Engineering*.

Koay, G., Kearns, D., Heffner, H. E., and Heffner, R. S. (1998). Passive sound-localization ability of the big brown bat (eptesicus fuscus). *Hearing Research*, 119(1-2):37–48.

Koka, K., Read, H. L., and Tollin, D. J. (2008). The acoustical cues to sound location in the rat: Measurements of directional transfer functions. *Journal of the Acoustical Society of America*, 123(6):4297–4309.

Kössl, M. and Russell, I. J. (1995). Basilar-membrane resonance in the cochlea of the moustached bat. *Proceedings of the National Academy of Sciences of the United States of America*, 92(1):276–279.

Kössl, M. and Vater, M. (1995). *Hearing by bats: Cochlea structure and function in bats, in A. N. Popper and R. R.R. Fay, eds*, volume 5. New York: Springer-Verlag.

Kuc, R. (1996). Biologically motivated adaptive sonar system. *Journal of the Acoustical Society of America*, 100(3):1849–1854.

Kuc, R. (1997a). Biomimetic sonar locates and recognizes objects. *Ieee Journal of Oceanic Engineering*, 22(4):616–624.

Kuc, R. (1997b). Biomimetic sonar recognizes objects using binaural information. *Journal of the Acoustical Society of America*, 102(2):689–696.

Kuc, R. (2002). Binaural sonar electronic travel aid provides vibrotactile cues for landmark, reflector motion and surface texture classification. *Ieee Transactions on Biomedical Engineering*, 49(10):1173–1180.

Kuc, R. (2007). Biomimetic sonar and neuromorphic processing eliminate reverberation artifacts. *Ieee Sensors Journal*, 7(3-4):361–369.

Lawrence, B. D. and Simmons, J. A. (1982a). Echolocation in bats - the external ear and perception of the vertical positions of targets. *Science*, 218(4571):481–483.

Lawrence, B. D. and Simmons, J. A. (1982b). Measurements of atmospheric attenuation at ultrasonic frequencies and the significance for echolocation by bats. *Journal of the Acoustical Society of America*, 71(3):585–590.

Liu, J., Erwin, H., and Wermter, S. (2008). Mobile robot broadband sound localisation using a biologically inspired spiking neural network. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2191–2196.

Lyon, R. F. (1996). The all-pole gammatone filter and auditory models. *Acustica*, 82:S90.

Maki, K. and Furukawa, S. (2005). Acoustical cues for sound localization by the mongolian gerbil, meriones unguiculatus. *Journal of the Acoustical Society of America*, 118(2):872–886.

Masters, W. M., Raver, K. A. S., and Kazial, K. A. (1995). Sonar signals of big brown bats, eptesicus-fuscus, contain information about individual identity, age and family affiliation. *Animal Behaviour*, 50:1243–1260.

Middlebrooks, J. C. and Green, D. M. (1990). Directional dependence of interaural envelope delays. *Journal of the Acoustical Society of America*, 87(5):2149–2162.

Mogdans, J., Ostwald, J., and Schnitzler, H. U. (1988). The role of pinna movement for the localization of vertical and horizontal wire obstacles in the greater horseshoe bat, rhinolopus-ferrumequinum. *Journal of the Acoustical Society of America*, 84(5):1676–1679.

Moss, C. F., Bohn, K., Gilkenson, H., and Surlykke, A. (2006). Active listening for spatial orientation in a complex auditory scene. *Plos Biology*, 4(4):615–626.

Müller, R. (2004). A numerical study of the role of the tragus in the big brown bat. *Journal of the Acoustical Society of America*, 116(6):3701–3712.

Müller, R. and Kuc, R. (2007). Biosonar-inspired technology: goals, challenges and insights. *Bioinspiration and Biomimetics*, 2:S146–S161.

Müller, R. and Schnitzler, H. U. (2000). Acoustic flow perception in cf-bats: Extraction of parameters. *Journal of the Acoustical Society of America*, 108(3):1298–1307.

Neuweiler, G. (1990). Auditory adaptations for prey capture in echolocating bats. *Physiological Reviews*, 70(3):615–641.

Obrist, M. K., Fenton, M. B., Eger, J. L., and Schlegel, P. A. (1993). What ears do for bats - a comparative-study of pinna sound pressure transformation in chiroptera. *Journal of Experimental Biology*, 180:119–152.

Papadopoulos, T. and Allen, R. (2007). Experimental method for the acoustical modelling of the echolocation process in bats. *Proceedings of the Institute of Acoustics*, 29(3).

Patterson, R. D., Robinson, K., Holdsworth, J., Mckeown, D., Zhang, C., and Allerhand, M. H. (1992). *Complex sounds and auditory images, In Auditory Physiology and Perception, (Eds.) Y. Cazals, L. Demany, K. Horner*. Pergamon, Oxford.

Peremans, H. and Hallam, J. (1998). The spectrogram correlation and transformation receiver, revisited. *Journal of the Acoustical Society of America*, 104(2):1101–1110.

Rabinowitz, W. M., Maxwell, J., Shao, Y., and Wei, M. (1993). Sound localization cues for a magnified head: Implications from sound diffraction about a rigid sphere. *Presence*, 2(2):125–129.

Rosenblum, L. D., Gordon, M. S., and Jarquin, L. (2000). Echolocating distance by moving and stationary listeners. *Ecological Psychology*, 12(3):181–206.

Russell, I. J., Drexl, M., Foeller, E., Vater, M., and Kossl, M. (2003). Synchronization of a nonlinear oscillator: Processing the cf component of the echo-response signal in the cochlea of the mustached bat. *Journal of Neuroscience*, 23(29):9508–9518.

Saillant, P. A., Simmons, J. A., and Dear, S. P. (1993). A computational model of echo processing and acoustic imaging in frequency-modulated echolocating bats - the spectrogram correlation and transformation receiver. *Journal of the Acoustical Society of America*, 94(5):2691–2712.

Schnitzler, H. U., Moss, C. F., and Denzinger, A. (2003). From spatial orientation to food acquisition in echolocating bats. *Trends in Ecology and Evolution*, 18(8):386–394.

Schnitzler, H. U. and O. W. Henson, J. (1980). *Performance of airborne animal sonar system. Microchiroptera in Animal Sonar Systems edited by R. -G. Busnel and J. F. Fish*. Plenum, New York.

Shi, R. Z. and Horiuchi, T. K. (2007). A neuromorphic vlsi model of bat interaural level difference processing for azimuthal echolocation. *Ieee Transactions on Circuits and Systems I-Regular Papers*, 54(1):74–88.

Siemers, B. M. and Schnitzler, H. U. (2004). Echolocation signals reflect niche differentiation in five sympatric congeneric bat species. *Nature*, 429(6992):657–661.

Simmons, J. A. and Chen, L. (1989). The acoustic basis for target discrimination by fm echolocating bats. *Journal of the Acoustical Society of America*, 86(4):1333–1350.

Simmons, J. A., Ferragamo, M., Moss, C. F., Stevenson, S. B., and Altes, R. A. (1990). Discrimination of jittered sonar echoes by the echolocating bat, eptesicus-fuscus - the shape of target images in echolocation. *Journal of Comparative Physiology A-Sensory Neural and Behavioral Physiology*, 167(5):589–616.

Simmons, J. A., Kick, S. A., Lawrence, B. D., Hale, C., Bard, C., and Escudie, B. (1983). Acuity of horizontal angle discrimination by the echolocating bat, eptesicus-fuscus. *Journal of Comparative Physiology*, 153(3):321–330.

Simmons, J. A., Saillant, P. A., Wotton, J. M., Haresign, T., Ferragamo, M. J., and Moss, C. F. (1995). Composition of biosonar images for target recognition by echolocating bats. *Neural Networks*, 8(7-8):1239–1261.

Simmons, J. A. and Stein, R. A. (1980). Acoustic imaging in bat sonar - echolocation signals and the evolution of echolocation. *Journal of Comparative Physiology*, 135(1):61–84.

Simon, R., Holderied, M. W., and von Helversen, O. (2006). Size discrimination of hollow hemispheres by echolocation in a nectar feeding bat. *Journal of Experimental Biology*, 209(18):3599–3609.

Slaney, M. (1993). An efficient implementation of the patterson-holdsworth auditory filter bank. *Apple Computer Technical Report 35*.

Suga, N. and Jen, P. H. S. (1975). Peripheral control of acoustic signals in the auditory system of echolocating bats. *Journal of Experimental Biology*, 62(2):277–311.

Surlykke, A. and Kalko, E. K. V. (2008). Echolocating bats cry out loud to detect their prey. *PLoS ONE*, 3(4):e2036.

Teeling, E. C., Springer, M. S., Madsen, O., Bates, P., O'Brien, S. J., and Murphy, W. J. (2005). A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science*, 307(5709):580–584.

Ulanovsky, N. and Moss, C. F. (2008). What the bat's voice tells the bat's brain. *Proceedings of the National Academy of Sciences of the United States of America*, 105(25):8491–8498.

Vonderemde, G. and Menne, D. (1989). Discrimination of insect wingbeat-frequencies by the bat rhinolophus-ferrumequinum. *Journal of Comparative Physiology A-Sensory Neural and Behavioral Physiology*, 164(5):663–671.

Walker, V. A., Peremans, H., and Hallam, J. C. T. (1998). One tone, two ears, three dimensions: A robotic investigation of pinnae movements used by rhinolophid and hipposiderid bats. *Journal of the Acoustical Society of America*, 104(1):569–579.

Waters, D. and Abulula, H. (2001). The virtual bat: echolocation in virtual reality. *Proceedings of the 2001 International Conference on Auditory Display. Espoo, Finland*, pages 191–196.

Waters, D. A. (2007). Echolocation in air: biological systems, technical challenges, and transducer design. *Proceedings of the Institution of Mechanical Engineers Part C-Journal of Mechanical Engineering Science*, 221(10):1165–1175.

Waters, D. A. and Jones, G. (1995). Echolocation call structure and intensity in 5 species of insectivorous bats. *Journal of Experimental Biology*, 198(2):475–489.

Wiegrebe, L. (2008). An autocorrelation model of bat sonar. *Biological Cybernetics*, 98(6):587–595.