

受聴者の頭部形状に基づいた
頭部伝達関数の個人化に関する研究

A study on personalization of head-related transfer functions
based on the anthropometry of the listener's head

石井 要次

平成 30 年 3 月

要旨

本論文は、3次元音像制御精度の向上を図るため、受聴者の頭部形状に基づいて頭部伝達関数を個人化する方法を提案し、その有効性を示したものである。

頭部伝達関数に含まれるヒトの方向知覚の手掛かりは、前後・左右方向で異なる。左右方向知覚の手掛かりは両耳間差キューであり、両耳間差キューには両耳間時間差と両耳間レベル差があることが知られている。前後・上下方向知覚の手掛かりは、頭部伝達関数の第1、第2ノッチがスペクトラルキューとして重要な役割を果たしていることが知られている。頭部伝達関数の第1ピークと第1ノッチは耳介によって生成されることが示されている。また、第1、第2ノッチと第1ピークで再構成されたパラメトリック頭部伝達関数は、実測した頭部伝達関数と同等の定位精度が得られることが報告されている。ただし、上方では、一部の被験者においてパラメトリック頭部伝達関数の定位精度が低下することが報告されている。

頭部伝達関数には方向知覚の手掛かりが含まれているため、受聴者本人の頭部伝達関数を鼓膜上で再現することで任意の3次元方向に音像を制御することができる。しかし、頭部伝達関数には個人差があるため、他人の頭部伝達関数を用いると誤った方向に定位する現象がしばしば発生する。頭部伝達関数の個人差は、頭部や耳介形状の個人差に起因していることが報告されている。

3次元音像制御を一般に普及させるためには、各受聴者に適合する頭部伝達関数を容易に提供する方法が必要となる。試聴により適合する頭部伝達関数を選出する方法、および受聴者の頭部形状から適合する頭部伝達関数を求める方法が提案されているが、いずれの従来法においても、高い音像定位精度は得られていない。

本研究では、方向知覚の手掛かりに着目し、受聴者の頭部形状から方向知覚の手掛かりである両耳間差キューおよびスペクトラルキューを推定することで各受聴者に適合する頭部伝達関数を提供する方法を提案し、その有効性について検証した。また、上方の音像定位精度の向上を目的とし、頭部伝達関数の第2ピークが上方の音像定位精度に及ぼす影響を検証した。

本論文は以下の5つの章から構成されている。

第1章では、ヒトの方向知覚の手掛かり、頭部伝達関数について解説した上で、3次元音像制御における問題点として頭部伝達関数の個人差を採り上げた。また、本研究の目的として、受聴者の頭部形状から方向知覚の手掛かりである両耳間差キューおよびス

スペクトラルキューを推定することで各受聴者に適合する頭部伝達関数を提供する方法を提案し、その有効性について物理的側面および知覚的側面から検証することを述べた。

第2章では、各受聴者の頭部形状から左右方向知覚の手掛かりである両耳間時間差および両耳間レベル差を個人化する方法を提案し、従来法および方向知覚の弁別閾と比較することで提案法の有効性について検証した。

提案法では、受聴者の頭部形状の前後・左右非対称性を反映できる頭部モデルを導入した。33名の被験者の10箇所の頭部形状、水平面内12方向の両耳間時間差および両耳間レベル差を計測した。さらに、方位角ごとに、水平面内12方向の両耳間時間差、および1/3 octave bandの両耳間レベル差を目的変数、頭部形状を説明変数とした重回帰式を構築した。次に、4名のナイーブな被験者を用いて提案法の有効性を検証した。その結果、4名の被験者の平均推定誤差は、両耳間時間差ではいずれの方向においても弁別閾と同等であり、両耳間レベル差では800 Hz以下の帯域では弁別閾内であった。しかし、1000 Hz以上の両耳間レベル差の推定誤差は、弁別閾を超えた。

第3章では、受聴者の耳介形状から前後・上下方向知覚の手掛かりであるスペクトラルキューの個人化方法を提案し、提案法の有効性について検証した。本研究では、受聴者の耳介形状から前後・上下方向知覚の手掛かりとして重要な役割を果たす第1、第2ノッチの周波数を推定し、その周波数に近い第1、第2ノッチの周波数を持つ頭部伝達関数をデータベースから選出することで頭部伝達関数の個人化を行う方法を提案した。次に、提案法の有効性を検証するため、4名のナイーブな被験者の個人化した頭部伝達関数を用いて正中面内を目標方向とした音像定位実験を行った。その結果、上昇角0、180°においては実測した頭部伝達関数と同等の精度で音像制御を提供できることを示した。しかし、一部の被験者において上方で音像定位精度が低下する場合があった。

第4章では、頭部伝達関数の第2ピークを加えることで上方の音像定位精度が向上するか否か、および第2ピークが上方の音像定位精度に及ぼす影響について検証することを目的とし、第1、第2ノッチおよび第1ピークで再構成したパラメトリック頭部伝達関数と、それに第2ピークを加えたパラメトリック頭部伝達関数を用いて上半球正中面内を目標方向とした音像定位実験を行った。その結果、第1、第2ノッチおよび第1ピークで再構成したパラメトリック頭部伝達関数に第2ピークを付加することで上方の音像定位精度が向上し、実測した頭部伝達関数と有意な差がない音像定位精度となることを示した。また、第2ピーク自身はスペクトラルキューではないが、第2ピークは

第 1 ノッチを強調することにより, 正中面上方の音像定位に重要な役割を果たしている可能性を示した. 上半球正中面内全体において, 実測した頭部伝達関数と同等の定位精度を提供できる最小構成は, 第 1, 第 2 ノッチ, 第 1, 第 2 ピークであることが示唆された.

第 5 章では, 今後の展望等を踏まえて本論文を総括した.

Abstract

The present study proposes a method for the personalization of the head-related transfer function on the basis of the anthropometry of the listener's head.

Interaural time difference and interaural level difference are well-known horizontal localization cues, and the two lowest spectral notches play an important role in vertical localization. The parametric head-related transfer function recomposed of only the first peak and the two lowest spectral notches, which are respectively generated by the first resonance and the first and second anti-resonances of the pinna, has been reported to provide approximately the same localization performance as the measured HRTF does for the front and rear directions. However, for the upper direction, the localization performance for some of the subjects decreased.

HRTFs provide accurate 3D sound image localization. However, the HRTFs of other listeners often degrade localization accuracy. Individual differences in HRTF have been reported to be caused by individual differences in the anthropometry of the head and pinna.

For disseminating 3D sound image control, a method to provide the appropriate HRTF for each listener is necessary. A method to select the appropriate HRTF through a listening test and to estimate the appropriate HRTF by anthropometry of the listener's head has been proposed. The previous method could not obtain a high-accuracy sound image localization.

To improve the precision of 3D sound image production, this study proposes a method that provides the appropriate HRTFs for each listener. This is achieved by estimating sound image localization cues, which are the interaural difference cue and the spectral cue, on the basis of the listener's anthropometry of the head.

This paper is composed of the following five Chapters:

In Chapter 1, localization cues, the definition of HRTF, and individual differences in HRTFs as an issue in 3D sound image control are explained. The purpose of the study is also presented.

In Chapter 2, a method to estimate the interaural time difference and interaural level difference based on the anthropometry of the listener's head is studied. Ten anthropometric parameters of the heads, as well as the interaural time difference and interaural level difference on a horizontal plane for 33 subjects, were measured. Then, multiple regression analyses were carried out as objective variables of the interaural time difference and interaural level difference and as explanatory variables of ten anthropometric parameters of the head. The accuracy of the multiple regressions was verified using four naive subjects.

The mean estimation errors of the interaural time difference and interaural level difference, which for lower than 800 Hz, were almost the same as that of just noticeable difference. However, for those bands higher than 1,000 Hz, the estimation error was larger than that of just noticeable difference. Moreover, the estimation accuracy of the proposed method was compared with that of the previous method. The average estimation error of the proposed method was smaller than that of the previous method.

In Chapter 3, a method to estimate the appropriate HRTFs for an individual listener is studied. The proposed method estimates the frequencies of the two lowest spectral notches, which play an important role in vertical localization, in the HRTF of an individual listener by anthropometry of the listener's pinnae. The best-matching HRTFs, among which N1 and N2 are the closest to the estimates, are then selected from an HRTF database. Localization tests in the upper median plane were performed using four subjects to examine the validity of the proposed method. The results revealed that the best-matching HRTFs provided approximately the same performance as the listener's own HRTFs did for the target directions of the front and rear for all four subjects. For the upper target directions, however, the performance of the localization for some of the subjects decreased.

In Chapter 4, two localization tests with four listeners and seven target angles in the upper median plane were conducted to investigate whether adding a second spectral peak can resolve this performance decrease, and to examine the effects of the second spectral peak on the localization of the upper direction in the median

plane. The results suggested that adding the second spectral peak to the first and second notches and the first peak resulted in a decrease in the mean vertical localization error, and no statistically significant difference was observed between the added second peak and first and second notches and the measured HRTFs at any target vertical angle. A sound image was hardly perceived in the upper direction by reproducing only the second spectral peak, but the presence of the second spectral peak to improve the salience of the first notches was discussed. These findings that the minimum components, which provide approximately the same localization performance as the subject's own HRTF, are N1, N2, P1 and P2.

In Chapter 5, the conclusions are presented, and issues that can be investigated in the future are described.

目次

1	序論	1
1.1	本論文で扱う座標系	1
1.2	方向知覚の手掛かり	1
1.2.1	頭部伝達関数の定義	3
1.2.2	頭部伝達関数における方向知覚の手掛かり	4
1.3	3次元音像制御方式	10
1.3.1	耳入力信号再現法	10
1.3.2	波面合成法	12
1.3.3	境界音場制御法	12
1.3.4	アンビソニックス	13
1.3.5	3次元音像制御方式のまとめ	13
1.4	頭部伝達関数の個人差	14
1.5	受聴者に適合する頭部伝達関数の提供方法	16
1.5.1	頭部伝達関数の標準化	17
1.5.2	頭部伝達関数の個人化	17
1.6	3次元音像制御に必要な頭部伝達関数の情報	20
1.7	本論文で取り扱う研究課題	21
1.8	本研究の目的	21
2	両耳間差キューの個人化	24
2.1	本章の目的	24
2.2	両耳間時間差の個人化	24
2.2.1	従来の研究	24
2.2.2	本節の目的	25
2.2.3	両耳間時間差の算出	25
2.2.4	頭部寸法の特徴量の計測	28
2.2.5	単回帰分析	30
2.2.6	重回帰分析	31
2.2.7	ナイーブな被験者による両耳間時間差の推定精度の検証	34

2.3	両耳間レベル差の個人化	37
2.3.1	本節の目的	37
2.3.2	従来の研究	37
2.3.3	両耳間レベル差の算出	38
2.3.4	頭部形状の計測	42
2.3.5	単回帰分析	42
2.3.6	重回帰分析	45
2.3.7	ナイーブな被験者による両耳間レベル差の推定精度の検証	50
2.4	結論	51
3	スペクトラルキューの個人化	54
3.1	本章の目的	54
3.2	従来の研究	54
3.3	頭部伝達関数の測定	56
3.4	頭部伝達関数の第1第2ノッチの抽出	56
3.5	耳介形状の計測	57
3.6	耳介形状による第1第2ノッチ周波数の推定	59
3.6.1	重回帰分析	59
3.6.2	重回帰モデルの精度	59
3.7	ナイーブな被験者の第1, 第2ノッチ周波数の推定精度	60
3.7.1	物理的精度	60
3.7.2	心理的評価	66
3.8	考察	75
3.9	結論	76
4	頭部伝達関数の第2ピークが正中面上方の音像定位精度に及ぼす影響	78
4.1	本章の目的	78
4.2	パラメトリック頭部伝達関数の再構成方法	79
4.2.1	頭部伝達関数の測定	79
4.2.2	ノッチとピークの抽出	79
4.2.3	パラメトリック頭部伝達関数の生成	80
4.3	音像定位実験	81

4.3.1	音像定位実験(1)	81
4.3.2	音像定位実験(2)	86
4.4	考察	89
4.5	結論	90
5	総括	91
5.1	結論	91
5.2	今後の展望	93

目次

1.1 球座標系と矢状面座標系	2
1.2 音の伝搬経路	3
1.3 頭部伝達関数の一例	4
1.4 cone of confusion	5
1.5 正中面内の HRTF の振幅スペクトルと N1,N2 周波数	6
1.6 正中面における HRTF のピークノッチパターン	7
1.7 耳介部位の名称	7
1.8 N1 周波数による励振シミュレーション	9
1.9 波面合成法	12
1.10 境界音場制御法	13
1.11 高次アンビソニックス	13
1.12 本人および他人の HRTF を用いた 水平面, 正中面での音像定位実験の結果	15
1.13 正面方向の HRTF	16
1.14 耳介形状	16
1.15 受聴者に適合する HRTF の提供方法	17
1.16 本研究における HRTF の個人化方法のプロセス	23
1.17 本論文の章構成	23
2.1 hrir の測定機器の系統図	26
2.2 全被験者の ITD の方位角に対する変化	28
2.3 ITD の推定に用いた頭部寸法の計測箇所	29
2.4 33 名の水平面での ILD	40
2.5 2 名の水平面での ILD の比較	41
2.6 ILD の推定に用いた頭部モデル	42
2.7 提案法の重回帰モデルによる推定誤差の絶対値	51
2.8 周波数帯域ごとの平均推定誤差	52
3.1 耳介形状の計測箇所	57
3.2 被験者本人の HRTF と best-matching HRTF の振幅スペクトル	63

3.3	正面方向の本人 HRTF と best-matching HRTF の N1,N2 周波数の散布図	65
3.4	マッピング法による回答画面	67
3.5	実音源, 本人の HRTF, best-matchingHRTF による 正中面音像定位実験の結果	69
3.6	OIS の HRTF に対する各被験者の回答分布	73
3.7	TCY の HRTF に対する各被験者の回答分布	73
4.1	上昇角 90°のパラメトリック HRTF(N1N2P1, N1N2P1+P2)の例	81
4.2	実験結果 : 回答の分布	83
4.3	上昇角 90°における実測 HRTF と パラメトリック HRTF(P1, P2, P1P2)の一例	86
4.4	回答の分布	87
4.5	平均上昇角誤差	88
4.6	N1, N2, P1, P1 の周波数とレベルの関係	89

表目次

2.1	33名のITDの統計値	28
2.2	計測箇所とその名称	29
2.3	33名の頭部形状の統計値	30
2.4	頭部形状と各方向のITDの単相関係数	31
2.5	重回帰分析の結果. 重相関係数, 危険率, 残差の絶対値の平均, 定数項	33
2.6	ナイーブな被験者のITD	35
2.7	ナイーブな被験者の頭部寸法	35
2.8	提案法の重回帰モデルによる推定誤差の絶対値	35
2.9	Algazi <i>et al.</i> の重回帰モデルによる推定誤差の絶対値	35
2.10	提案法の重回帰モデルによる方位角誤差の絶対値	35
2.11	Algazi <i>et al.</i> の重回帰モデルによる方位角誤差の絶対値	35
2.12	ILDの左右・前後非対称性の平均値と最大値	41
2.13	頭部形状と各方向のILDの単相関係数 r (250 Hz)	44
2.14	頭部形状と各方向のILDの単相関係数 r (1000 Hz)	44
2.15	頭部形状と各方向のILDの単相関係数 r (4000 Hz)	44
2.16	重回帰分析の結果: 重相関係数	46
2.17	重回帰分析の結果: 重回帰係数の全帯域の平均値	47
2.18	重回帰分析の結果: 危険率	48
2.19	重回帰分析の結果: 残差の絶対値の平均	49
3.1	耳介形状の計測結果	58
3.2	重回帰係数, p 値, 95%信頼区間	60
3.3	重回帰モデルの統計値	60
3.4	4名のナイーブな被験者の耳介形状	61
3.5	4名の被験者のN1,N2周波数の推定誤差	61
3.6	正面方向における本人HRTFとbest-matching HRTFの N1,N2周波数	64
3.7	正中面7方向における被験者本人のHRTFとbest-matching HRTFの	

N1, N2 周波数の差	65
3.8 平均定位誤差	70
3.9 前後誤判定率	71
3.10 カイ二乗検定の結果	72
3.11 OIS と TCY の HRTF, best-matching HRTF の平均定位誤差	74
3.12 OIS と TCY の HRTF, best-matching HRTF の前後誤判定率	75
3.13 非個人化 HRTF と best-matching HRTF の前後誤判定率に対する カイ二乗検定の結果	75
4.1 平均上昇角定位誤差	84
4.2 Tukey の多重比較検定の結果	84
4.3 前後誤判定率	85
4.4 平均上昇角誤差に対する Tukey の多重比較検定の結果	88

第 1 章 序論

2016 年は VR(Virtual Reality)元年と呼ばれ、近年は VR 技術の普及が急速に進み、特に、映像に関しては、3D 映像や 360°映像等の高い臨場感が得られる技術が普及し、一般のユーザも容易に VR を体感できる。

一方、音については以下に示す音像の性質[1]を精度高く再現することができれば、ユーザに高い臨場感を持つ音空間を提供できる。

- ・ 時間的性質 残響感, リズム感, 持続感など
- ・ 空間的性質 方向感, 距離感, 広がり感など
- ・ 質的性質 大きさ, 高さ, 音色など

時間的性質と質的性質については、近年の音データの収録・再生技術の進歩、および容量の増加により、ハイレゾリューションオーディオのサンプリング周波数 386 kHz、量子化ビット数 36 bit や、direct stream digital のサンプリング周波数 11.2 MHz に対応した製品が市販される等、高い精度での再現が可能となっている。一方で、空間的性質における方向感ハステレオ再生や 5.1 ch サラウンドシステムが主流であり、立体的な音環境を一般の受聴者に提供するまでには至っていない。また、空間的性質の広がり感や距離感の制御方法についても一般に普及していない。

精度の高い 3 次元音像制御が実用化されれば、実在する音空間の時間や空間を超えた忠実な再現や、実在しない音空間のシミュレーション等を提供することができる。3 次元音像制御技術は、視覚障害者に対する補助等に応用することも考えられており、バリアフリー社会に対して貢献することが期待できる。立体的な音環境を再現、あるいは生成する技術として、耳入力信号再現法、波面合成法、アンビソニックス等の技術が提案されているが、いずれの方法も一般に普及していない。

本論文では、ヘッドホンや 2 個のスピーカによって 3 次元音像制御が実現できる耳入力信号再現法の普及の阻害要因の 1 つとなっている頭部伝達関数の個人差の問題を採り上げた。他人の頭部伝達関数を用いると、誤った方向に音像が定位するなど、音像制御精度が低下する現象がしばしば発生する。頭部伝達関数を測定するためには、特別な設備や多大な時間が必要であるため、全ての受聴者の頭部伝達関数を測定することは現実的ではない。また、頭部伝達関数データベースから、受聴者が試聴して選出する方法ではデータベースの規模が大きくなるにつれて試聴に多大な時間を要する。3 次元音像

制御技術を一般に普及させるためには、受聴者の試聴や特殊な設備を必要としない簡易的な方法で、各受聴者に適合する頭部伝達関数を提供する方法を確立する必要がある。

本論文では、受聴者の頭部形状に基づいた頭部伝達関数の個人化方法を提案し、その有効性について物理的あるいは知覚的側面から検証した研究成果を述べる。

1.1 本論文で扱う座標系

まず、本論文で扱う座標系について定義する。本論文では、音源や音像の位置を定義するために2つの座標系を用いている[1]。図1.1に座標系を示す。図1.1(a)に示す球座標系は、音源や音像の方向を方位角 ϕ と仰角 θ によって定義する。方位角 ϕ と仰角 θ の範囲は、それぞれ $0^\circ \leq \phi < 360^\circ$ 、 $-90^\circ \leq \theta \leq 90^\circ$ である。一方、図1.1(b)に示す矢状面座標系は、音源や音像の方向を側方角 α と上昇角 β によって定義する。側方角 α と上昇角 β の範囲は、それぞれ $-90^\circ \leq \alpha \leq 90^\circ$ 、 $0^\circ \leq \beta < 360^\circ$ である。

また、正中面はヒトの頭部を左右に分かつ面、水平面は耳珠と眼窩点がつくる面、横断面は人の頭部を前後に分かつ面と定義する。

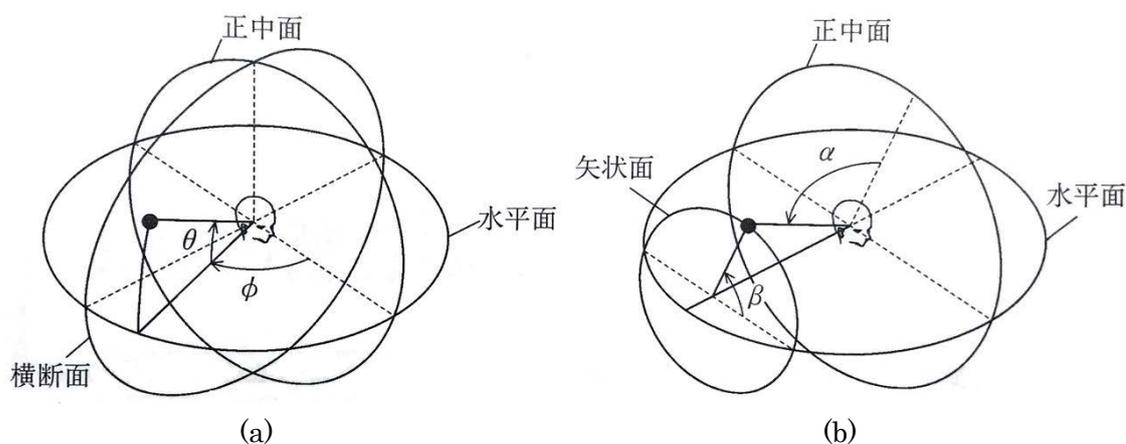


図 1.1 球座標系(a)と矢状面座標系(b) [1].

ϕ :方位角, θ :仰角, α :側方角, β :上昇角

1.2 方向知覚の手掛かり

ヒトの方向知覚に対して重要な役割を果たす物理特性の1つが頭部伝達関数(HRTF: Head-Related Transfer Function)である。まず、HRTFについて定義し、HRTFに含まれる方向知覚の手掛かりについて概説する。

1.2.1 頭部伝達関数の定義

まず、音源から放射された音が受聴者の鼓膜に到達するまでの伝達経路を考える。図 1.2 に示すように、音源から発せられた音は、室の反射や回折の時間特性を表す空間インパルス応答の影響を受けて受聴者の頭部近傍に到達する。さらに、受聴者の頭、耳介、胴体などによる時間的伝達特性を表す頭部インパルス応答(HRIR: Head-Related Impulse Response)の影響を受けて外耳道入口に到達し、外耳道を経て鼓膜に届く。HRTF は HRIR を周波数領域で表したものであり、式 1.1 で定義される。

$$HRTF_{l,r}(s, \alpha, \beta, d, \omega) = \frac{G_{l,r}(s, \alpha, \beta, d, \omega)}{F_0(\alpha, \beta, d_0, \omega)} \quad (1.1)$$

ここで、 $G_{l,r}(s, \alpha, \beta, d, \omega)$ は自由空間における音源から受聴者の外耳道入口もしくは鼓膜までの伝達関数、 $F_0(\alpha, \beta, d_0, \omega)$ は自由空間における受聴者がいない状態での音源から受聴者の頭部中心に相当する位置までの伝達関数である。添字の l, r は耳の左右、 s は受聴者、 α は音源の側方角、 β は音源の上昇角、 d は音源の距離、 ω は角周波数をそれぞれ表す。ただし、距離 d は 1 m を超えると HRTF には影響しない[2]。

頭部や耳介の形状は前後、上下で非対称であるため、HRTF は音の入射方向に依存して変化する。一方、外耳道は単純な筒状であるため、約 17 kHz 以下の周波数領域では 1 次元音場と見なせ、その伝達関数は音の入射方向によらない。このため、HRTF は外耳道入口で定義されることが多い[3]。

図 1.3 に正面方向の HRTF の一例を示す。0 dB は頭部がない状態での頭部中心に相当する位置での音圧レベルである。3–4 kHz に約 10 dB のピークがあり、6 kHz、8 kHz および 11 kHz 付近にそれぞれ鋭いノッチ（谷）がある。HRIR は約 2–3 ms で収束する。

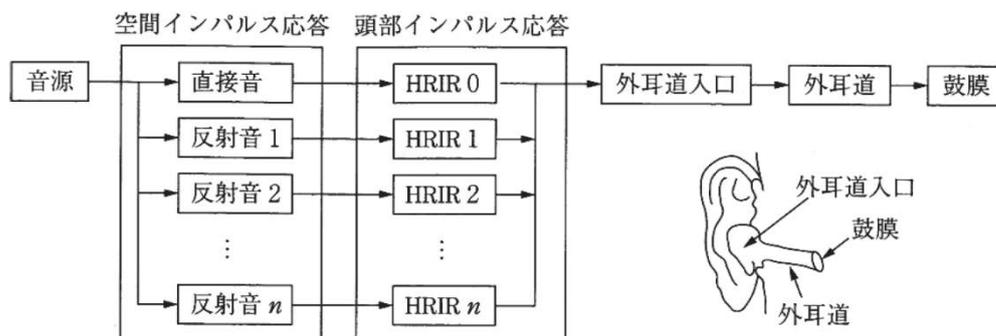


図 1.2 音の伝搬経路 [7]

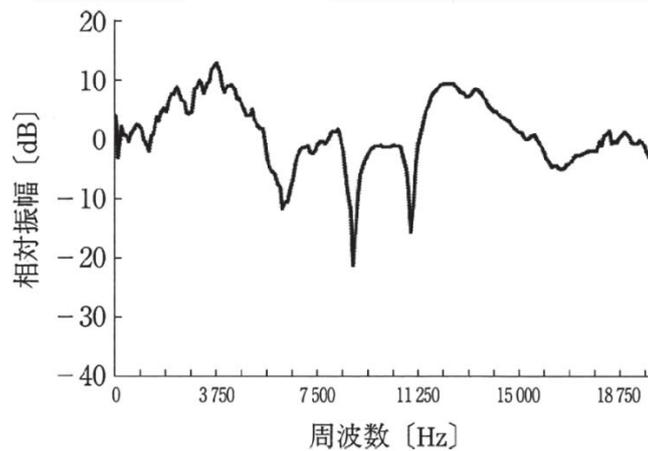


図 1.3 頭部伝達関数の一例 [7]

1.2.2 頭部伝達関数における方向知覚の手掛かり

HRTFに含まれるヒトの方向知覚の手掛かりは、左右方向と前後・上下方向で異なる。左右方向知覚の手掛かり、前後・上下方向知覚の手掛かりについて、それぞれ概説する。

A) 左右方向知覚の手掛かり

左右方向知覚の手掛かりは、両耳間時間差(ITD: interaural time difference)と両耳間レベル差(ILD: interaural level difference)であることが古くから知られている[4,5]。これらは両耳間差キューと呼ばれている。

左右の方向感は、ITD や ILD に対してほぼ線形に変化する。ITD が $0 \mu\text{s}$ で正面に知覚し、約 $1000 \mu\text{s}$ で側方に収束する[6]。両耳入力信号の波形そのものの時間差が左右方向の知覚の手掛かりとなるのは約 1600 Hz 以下の成分に限られる。これは、内耳の神経パルス発火に絶対不応期が存在するからである。約 1600 Hz 以上の周波数帯域では両耳入力信号の包絡線に対応した時間差が検出される[7]。

ILD については、 0 dB で正面に、約 10 dB で側方に知覚する[8]。ILD は可聴周波数帯域全域に亘って左右方向知覚の手掛かりとなっている。ただし、ILD は周波数依存性を持つため、入射方向が同じでも ILD は周波数により異なる。

ここで、頭部が球形であると仮定すると、図 1.4 に示す円錐台の垂直断面の円周上では ITD や ILD は等しくなる。したがって、両耳間差情報は音源が正中面から左右にどれだけ離れた面上にあるかを説明することができるが、前後・上下方向を同定することはできない。これはコーン状の混同(cone of confusion)と呼ばれている (図 1.4)。

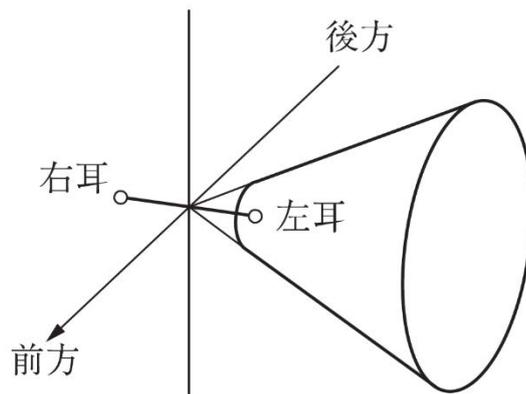


図 1.4 cone of confusion [7]

B) 前後上下方向知覚の手掛かり

前後・上下方向知覚の手掛かりは，HRTF の振幅スペクトルにあり，スペクトラルキューと呼ばれている．音像定位実験を用いた従来の多くの研究結果は，HRTF の 4 kHz 以上の振幅スペクトル情報が重要であるという点で一致している[9-12]．また，振幅スペクトルの微細構造はさほど重要でなく，むしろピークとノッチの概形が重要であることが明らかになっている[13,14]．

Iida *et al.* [15] は上昇角音像定位をシミュレートするパラメトリック HRTF モデルを提案した．パラメトリック HRTF は，音源の上昇角に依存しない 4 kHz のピーク[16]を下限周波数として，受聴者の実測 HRTF から抽出したノッチやピークから再構成される．ノッチとピークは周波数の順に「P1, N1, P2, N2, …」のようにラベルがつけられる．ノッチとピークは中心周波数，レベル，先鋭度でパラメトリックに表現される．彼らは上半球正中面において音像定位実験を行い以下のことを示した．

(1) すべてのノッチとピークで再構成されたパラメトリック HRTF は被験者本人の実測 HRTF と同等の定位性能を有する．

(2) 4 kHz 付近の第 1 ピーク(P1)と，P1 周波数より高く最も周波数の低い 2 つのノッチ(N1, N2)だけで再構成されたパラメトリック HRTF は，前方と後方においては被験者本人の実測 HRTF と同等の定位性能を有する．

(3) しかし上方においては，N1, N2, P1 で再構成されたパラメトリック HRTF の音像定位性能は，被験者によっては本人の実測 HRTF と比較して低下する．

(4) N1 と N2 の周波数は音源の上昇角に強く依存するが，P1 周波数はほぼ一定で上昇角に依存しない．

これらの結果より、彼らは N1 と N2 は、少なくとも前方と後方では、スペクトラルキューとして重要な役割を果たしていると結論付けた。さらに彼らは、ヒトの聴覚システムが耳入力信号から N1 と N2 を分析するための参照情報として P1 を用いている可能性を指摘した。図 1.5 に正中面内の HRTF の振幅スペクトルと N1,N2 周波数を示す。

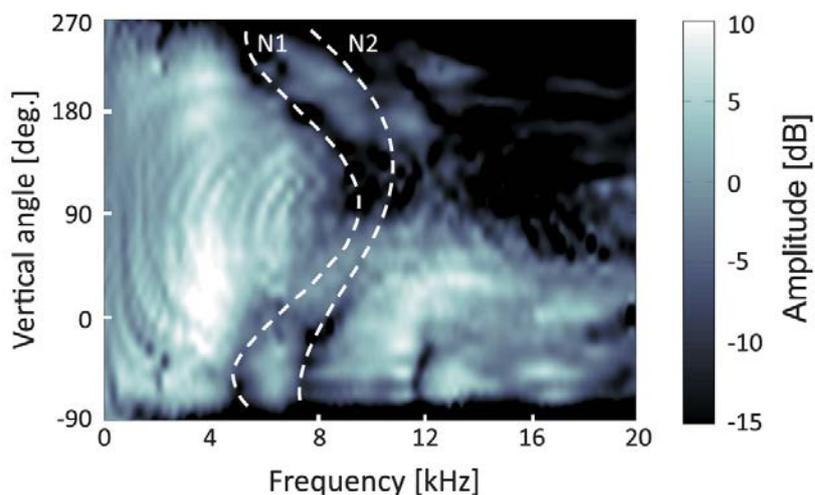


図 1.5 正中面内の HRTF の振幅スペクトルと N1,N2 周波数[15]

同様に、Zhang and Hartmann[17]は、前後知覚においては、ピークよりもノッチの影響が大きいことを報告しており、Moore *et al.* [18]はノッチ周波数の違いを受聴者は弁別できることを示している。つまり、音源の上昇角の変化に伴って HRTF のノッチ周波数が変化し、それを受聴者が弁別して前後・上下方向を知覚していると考えられる。

次に、スペクトラルキューの成因について考える。閉塞した外耳道入口で測定した正中面内の HRTF には、20 kHz までに 6 つのピークがあると言われている[19]。図 1.6 に周波数が低い方から 3 番目のピーク(P1, P2, P3)とノッチ(N1, N2, N3)を模式的に示す。横軸は周波数、縦軸は音源の上昇角を示しており、実線はピーク周波数、破線はノッチ周波数を示している。ピーク周波数は音源の仰角によらず一定であるのに対して、ノッチ周波数は音源が前方から上方に移動するにつれて高くなり、上方から後方に移動するにつれて低くなる。すなわち、ノッチ周波数の変化は音源の上昇角の変化と対応しているので、これらは音源の前後・上下方向知覚の手掛かりとなる。このようなピーク・ノッチの基本的なパターンは、耳介のみによって生じることが明らかになっている。耳介の各部位の解剖学的な名称を図 1.7 に示す。

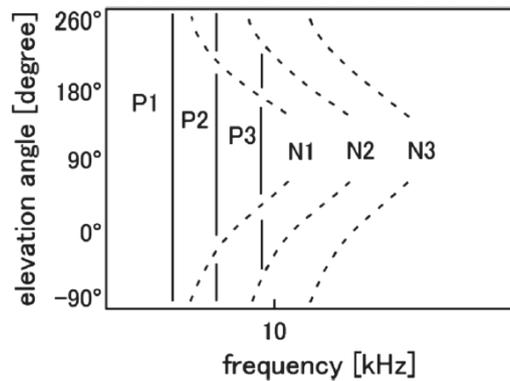


図 1.6 正中面における HRTF のピークノッチパターン[19]

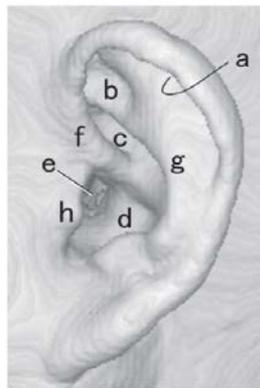


図 1.7 耳介部位の名称

a:舟状窩, b:三角窩, c:耳甲介舟, d:耳甲介腔,
e:外耳道, f:耳輪脚, g:対輪, h:耳珠[20]

ピークの成因は、耳介の共鳴モード（以降単にモードと呼ぶ）であることが、Shaw による実験[19]、Kahana and Nelson による境界要素法を用いたシミュレーション[21]、Takemoto *et al.*による時間領域差分法[20,22,23]を用いたシミュレーションなどで示されている。これらの研究では、ピーク周波数で励振した時に耳介の腔に生じる音圧の腹（以降単に腹と呼ぶ）の分布、強度、位相差などを分析している。その結果は定性的に共通しており、ピークは耳介に生じる様々なモードのうち、外耳道入口を含む耳甲介腔に腹が生じるものである。つまり、P1 では耳甲介腔に 1 つ、P2 では耳甲介腔に 1 つとその他の部位に 1 つ、P2 では耳甲介腔に 1 つとその他の部位に 2 つの腹が生じる。すなわち、ピークの番号が増加するにつれて腹の数も増加する。これは P4 以降でも同様であ

る. なお, P2 と P3 では, 2 つあるいは 3 つの腹が耳介の垂直方向に並ぶことから, vertical mode と呼ばれる.

P1 の成因は, 耳介の内外方向の第 1 次モード, すなわち, 耳甲介腔の深さを 1/4 波長とするモードである[19-21]. そのため, このモードでは耳介の腔全体にわたって 1 つの腹が生じる.

P2 の成因は, 耳介の上下方向の第 1 次モードである[19-21]. これは耳介の表面に沿って生じるモードであり, 外耳道入口の近傍と耳介の上部に互いに逆相の腹が 1 つずつ生じる. この耳介の上部に生じる腹の位置は研究によって異なっており, Shaw[19]と Takemoto *et al.* [20]は耳甲介舟と三角窩を合わせた部分に生じると報告し, Kahana and Nelson[21]は耳甲介舟と舟状窩を合わせた部分に生じると報告している. この違いは耳介形状の個人差を反映していると考えられるが, 耳介形状と腹が生じる位置の関係については解明されていない.

P3 の成因は, 耳介の上下方向の第 2 モードである[19-21]. このモードでは, 外耳道入口の近傍に 1 つとそれより上部の腔に 2 つの腹が生じる. 外耳道入口の腹に対し, 上部の 2 つの腹のうち, 外耳道入口に近い腹は逆相, 遠い腹は同相となる. P2 と同様に, これら耳介の上部に生じる 2 つの腹の位置は研究によって異なり, Shaw[19]は耳甲介舟と三角窩, Kahana and Nelson[21]は耳甲介舟と舟状窩, Takemoto *et al.* [20]は耳甲介腔の後方と三角窩に生じると報告している. この違いも耳介形状の個人差を反映しているものだと考えられるが, 腹の位置を決定する形態的な要因は見つかっていない.

P4 以上のピークの成因は, より複雑なモードであり, 詳細な検証は行われていない. ただし, P4 や P5 が前後・上下方向知覚に及ぼす影響は小さい.

ノッチの成因については, Raykar *et al.* [24]は, ノッチの成因となる節は, 外耳道入口に到達する直接波と, 耳介で反射して外耳道入口に到達する反射波とが経路差によって打ち消し合って生じるとして, 式 1.2 を示した.

$$f_n(\varphi) = \frac{(2n+1)}{2t_d(\varphi)}, \quad n=0,1,\dots \quad (1.2)$$

ここで, f_n はノッチ周波数[Hz], t_d は直接波と反射波の経路差による時間差[s], φ は直接波の入射角, n はノッチ番号である.

この式は, 音源が前方にある場合, 耳甲介腔の後壁などの上に反射点が見出せば適

用できる可能性がある。Raykar *et al.*は、ノッチから推定した反射点が対輪等の輪部に沿っていることを示した。すなわち、上昇角に対するノッチ周波数の変動は、耳介の形状を反映していることを示唆する。しかしこの式は、彼らが論文中でも述べているように、音源が後上方にある場合、反射点とする耳介部位が見出せないため適用できない。また、正面方向の音源に限ってもこの式で求めた周波数は受聴者のノッチ周波数と一致しないことが指摘されている[25]。

Takemoto *et al.* [20] は、N1 の成因に着目し、直接波と逆の位相を持つ腹が耳介上に生じることによってノッチの成因となる節が生じるとした。図 1.8 は時間領域差分法によって N1 周波数で耳介の腔を励振したときの瞬時音圧の分布を示している。音源が下方や前上方にあるとき (図 1.8a,b), 進行波の波面は、外耳道入口に対して反対側の耳介上に生じた逆相の腹によって打ち消され、外耳道入口付近に節が生じる。音源が後上方にあるとき (図 1.8c), 耳介の外側を通過する進行波の波面は耳介の内側にある耳甲介舟に生じた逆相の腹によって打ち消され、外耳道入口付近に節が生じる。以上のように、N1 の成因については検証が進められているが、N2 の成因に関する検証は行われていない。

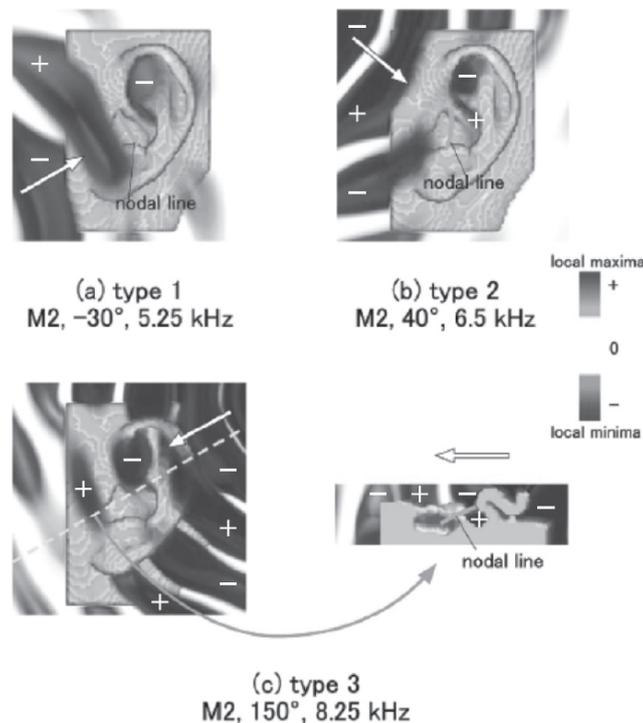


図 1.8 N1 周波数による励振シミュレーション[20]

1.3 3次元音像制御方式

従来提案されている3次元音像制御方式は、受聴者の耳入力信号を制御するもの、受聴者が聴取する位置の周辺の波面を制御するものに大別できる。

本節では、以下に示す代表的な3次元音像制御方式について概説する。

- 1) 耳入力信号再現法
- 2) 波面合成法
- 3) 境界要素法
- 4) アンビソニック

1.3.1 耳入力信号再現法

HRTFには方向知覚の手掛かりが含まれているため、HRTFを再現した音響信号 $p_{l,r}(t)$ を鼓膜に再現することで任意の方向に音像を制御することができる。

$$p_{l,r}(t) = F^{-1}\{S(\omega) \times HRTF_{l,r}(\omega, \theta, \phi)\} \quad (1.3)$$

ここで、 $S(\omega)$ は音源信号、添え字の l,r は左右の耳を表す。

これを実現する方式として、ヘッドホンを用いるバイノーラル方式、2個以上のスピーカを用いるトランスオーラル方式がある。以降、それぞれの原理について述べる。

a) バイノーラル再生方式

Møller *et al.* [26]は、ヘッドホン再生において、外耳道入口を開放した状態における鼓膜での音圧は、閉塞した状態の外耳道入口に以下の補正 G を施すことによって得られることを示している。

$$G = \frac{1}{M \cdot PTF} \cdot \frac{Z_{\text{ear canal}} + Z_{\text{headphone}}}{Z_{\text{ear canal}} + Z_{\text{radiation}}} = \frac{1}{M \cdot PTF} \cdot PDR \quad (1.4)$$

ここで、 M は耳栓型マイクロホンの伝達関数、 PTF は閉塞した外耳道入口で測定したヘッドホンの伝達関数、 $Z_{\text{ear canal}}$ と $Z_{\text{headphone}}$ はそれぞれ外耳道およびヘッドホンのインピーダンス、 $Z_{\text{radiation}}$ は外耳道から外側をみた放射インピーダンスである。 PDR はpressure division ratioの略であり、 PDR がほぼ1とみなせるFEC(free air equivalent coupling to the ear)ヘッドホンを利用すれば、ヘッドホンから外耳道を閉塞した状態での外耳道入口までの伝達関数の逆フィルタと受聴者のHRTFを畳込んだ刺激を提示することで、目標の方向に音像を制御することが可能である。Wightman and Kistler[27]は、プローブマイクロホンにより被験者の外耳道内の鼓膜の直前で測定したHRTFを

用いて、方位角が0–360°、仰角が-36–54°の範囲の72方向を目標方向としてヘッドホン実験を行った。その結果、HRTFを用いたヘッドホン再生によって、無響室における実音源を用いた場合と同等の定位精度が得られることを報告している。

b) トランスオーラル再生方式の原理

最小構成である2個のスピーカを用いた場合、左(右)のスピーカから左耳(右耳)に直接到達するダイレクト成分、およびの左(右)のスピーカから右耳(左耳)に到達するクロストーク成分が存在する。原理的には、以下の式によってクロストーク成分をキャンセルし、目標方向のHRTFを外耳道入口で再現すれば任意の方向に音像を制御可能である[28-31]。

$$\left. \begin{aligned} X_L(\omega) &= S(\omega) \times \frac{H_l(\omega) \times C_{R,r}(\omega) - H_r(\omega) \times C_{R,l}(\omega)}{C_{L,l}(\omega) \times C_{R,r}(\omega) - C_{L,r}(\omega) \times C_{R,l}(\omega)} \\ X_R(\omega) &= S(\omega) \times \frac{H_r(\omega) \times C_{L,l}(\omega) - H_l(\omega) \times C_{L,r}(\omega)}{C_{L,l}(\omega) \times C_{R,r}(\omega) - C_{L,r}(\omega) \times C_{R,l}(\omega)} \end{aligned} \right\} \quad (1.5)$$

ここで、 $X_i(\omega)$ はスピーカから放射される刺激、 $S(\omega)$ は音源信号、 $H_j(\omega)$ は目標方向に対する耳 j のHRTF、 $C_{ij}(\omega)$ はスピーカ i から耳 j への再生系伝達関数であり、 i は左右のスピーカ(L, R)、 j は左右の耳(l, r)を表す。

Morimoto and Ando[2]は、2個のスピーカによるトランスオーラル再生を用いて、正中面および水平面内を目標方向とした音像定位実験を行った。その結果、無響室における実音源と同等の定位精度が得られることを示している。

トランスオーラル(スピーカ再生)における2個のスピーカの配置については、水平面内の前方 $\pm 30^\circ$ が標準的であるが、 $\pm 5^\circ$ に近接配置するStereo Dipole[32]も提案されている。しかし、受聴時の位置のわずかなずれや頭部の動きにより、再生系伝達関数は必ずしも一定ではないため、実際には3次元音像制御は達成できているとは言えない。Morimoto and Ando[2]は、被験者の頭部を固定し、再生系伝達関数に顕著なスペクトラルノッチがみられないことから2個のスピーカを横断面の天頂から $\pm 30^\circ$ に配置したトランスオーラルシステムで音像定位実験を行った。その結果、受聴者本人のHRTFを使えば、実音源と同等の精度で音像を制御できることを示した。Takeuchi *et al.* [33]はStereo Dipoleを用いて、頭部を固定せずに水平面内の音像制御実験を行った。その結果、比較的高い精度で制御できた被験者のグループと、制御精度が低く水平面内の後方には定位しない被験者のグループに分かれたと報告している。頭部の固定は実用化の観点からは

許容し難い拘束条件であり、定められた位置に設置された椅子に座れば頭部を固定せずとも任意の3次元方向に音像を制御できることが望まれる。

制御領域を拡張するために、3個以上のスピーカを用いる方法も検討されているが[34, 35], 飯田らは、頭部を固定せずとも2個のスピーカを用いたトランスオーラルにおいて精度高く音像制御できるスピーカ配置を検討した。横断面15方向のスピーカ配置において、音像定位実験と再生系伝達関数の分析を行った結果、側方から上方にスピーカを配置した場合に音像制御精度が高くなることを示した。特に、天頂から $\pm 70^\circ$ に配置した場合は、水平面定位においても上半球正中面定位においても、実音源と有意差のない定位精度となることが示されている。

1.3.2 波面合成法

収録音場を複数のマイクロホンによって取り囲んで音響信号を収録し、それと同じ位置に配置したスピーカによって再生音場の全域で原音場を再現する波面合成法[36,37]が提案されている。この方法は、受聴者の移動に強いことが特徴であるが、再生音場において無響室や数多くのスピーカが必要であるなど、システムが大規模であるため、実用性に乏しい。

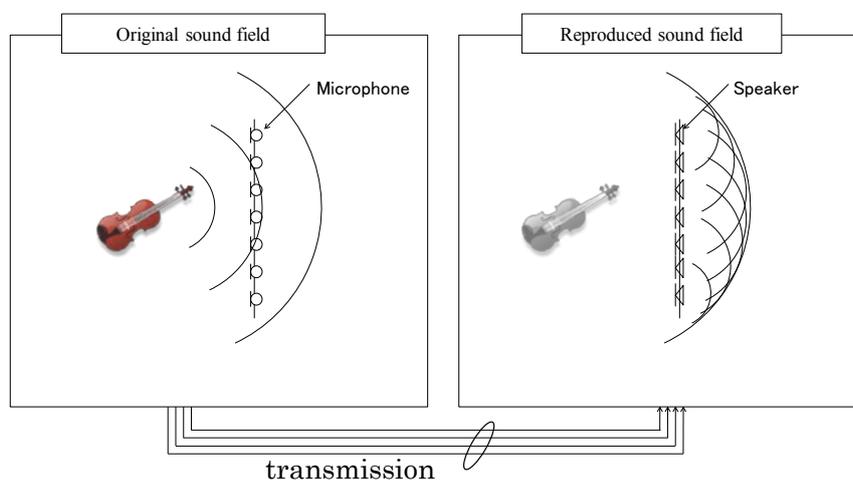


図 1.9 波面合成法

1.3.3 境界音場制御法

空間をいくつかの等立体角に分割した各小空間から到来する音響信号を指向性マイクロホンによって収録し、再生時は各小空間を代表するスピーカから再生することで、頭部中心位置付近で原音場の音響信号を再現する方法である[38]。頭部中心位置での音

響信号は忠実に再現できるが、外耳道入口での音響信号は再現することはできない。96個のスピーカを用いた音響樽等が提案されているが、大規模なシステムとなる[39].

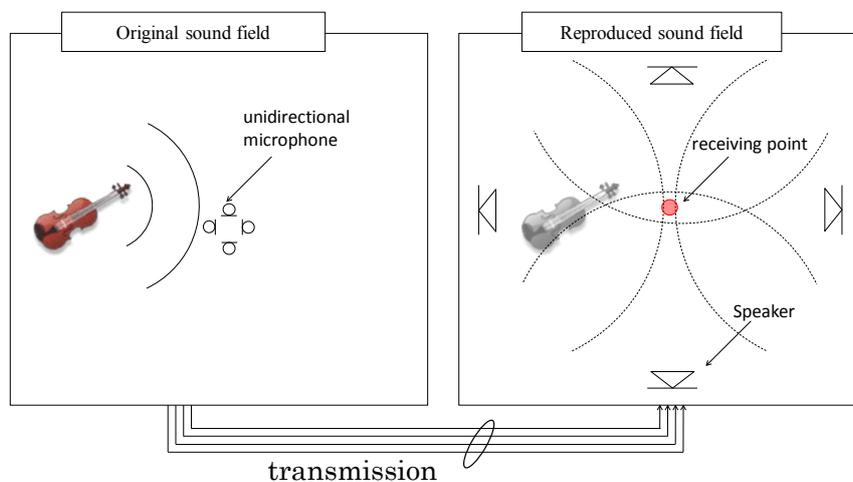


図 1.10 境界音場制御法

1.3.4 アンビソニックス

アンビソニックスは、指向性を持つ複数のマイクロホンによって収録し、0次と1次の球面調和関数を用いて音場を再現するものである[40]. 2次以上の球面調和関数を用いるものは高次アンビソニックスと呼ばれる[41]. この方法では、スピーカアレイの中心とその近傍において音場制御を行うことができる。157 chを用いた方法[42]等が提案されているが、境界音場制御法と同様に、大規模なシステムとなる。

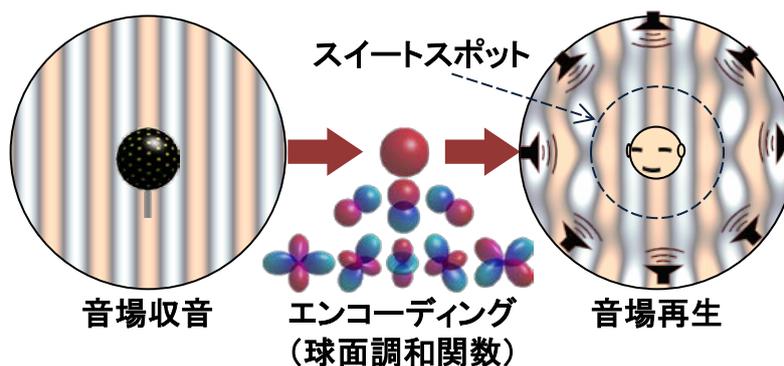


図 1.11 高次アンビソニックス[43]

1.3.5 3次元音像制御方式のまとめ

各方式の特徴をまとめる。波面合成法、境界音場制御方式、アンビソニックスについては、再生系において多くのスピーカ等が必要であり、大規模なシステムとなるため、

実用性が高いとは言い難い。一方、耳入力信号再現法は、ヘッドホンや2個のスピーカによって実現でき、大規模なシステムを必要としないことが実用性に関する利点の1つであると言える。

しかし、耳入力信号再現方式が一般に普及していない要因の1つに、受聴者の頭部形状の個人差に起因する HRTF の個人差が挙げられる。他人の HRTF を用いると、誤った方向に定位する現象がしばしば発生する。高い3次元音像制御を普及させるためには、各受聴者に適合する HRTF を提供する方法を確立する必要がある。次節では、HRTF の個人差について概説する。

1.4 頭部伝達関数の個人差

HRTF には受聴者の頭部形状や耳介形状の違いに起因する個人差が存在する。図 1.12 にトランスオーラル再生を用いて、本人および他人の HRTF を用いた水平面(a)および正中面(b)での音像定位実験の結果を示す[2]。各パネルの並びは、縦方向は被験者(L, M, S)を表し、横方向は使用した HRTF (被験者 L, M, S) を表す。一番右端のパネルは実音源を用いた場合の結果である。水平面(a)では、実音源、および本人の HRTF を用いた場合では概ね目標方向に定位している。他人の HRTF を用いた場合、前後誤判定や回答のばらつくなどの現象が発生している。一方、正中面(b)では、水平面と同様に、本人の HRTF を用いると概ね目標方向に定位しているが、他人の HRTF を用いた場合、音像の上昇や前後誤判定などの現象が発生している。なお、この実験ではトランスオーラル再生を用いているが、ヘッドホン再生を用いても同様の結果となることが示されている[44]。

次に、HRTF の個人差を物理的にみる。図 1.13 に、4名の左耳の正面方向の HRTF を示す。振幅スペクトルは、約 4 kHz 以下の周波数帯域では個人差は小さいが、それ以上の帯域ではピークやノッチの周波数やレベルの個人差が大きい。Yan *et al.* [45]は、国内外の5つの HRTF データベース[46-50]に含まれる 343名の正面方向の HRTF の N1, N2 周波数を求め、N1 周波数は 5063–12188 Hz(1.27 oct.)、N2 周波数は 7494–16882 Hz(1.17 oct.)にわたって広く分布することを報告している。ここで、N1, N2 周波数の JND は 0.1–0.2 oct. であることから[51]、N1, N2 周波数の個人差が音像定位精度に影響を及ぼしているといえる。また、Yan *et al.* [45]は、上記の5つのデータベースのうち、3つのデータベース間で 102名について図 1.14 に示す 9箇所(耳介形状)を比較した結果、各部

位において約 10 mm から 30 mm の個人差があることを報告している. これらの個人差が HRTF の個人差の要因である.

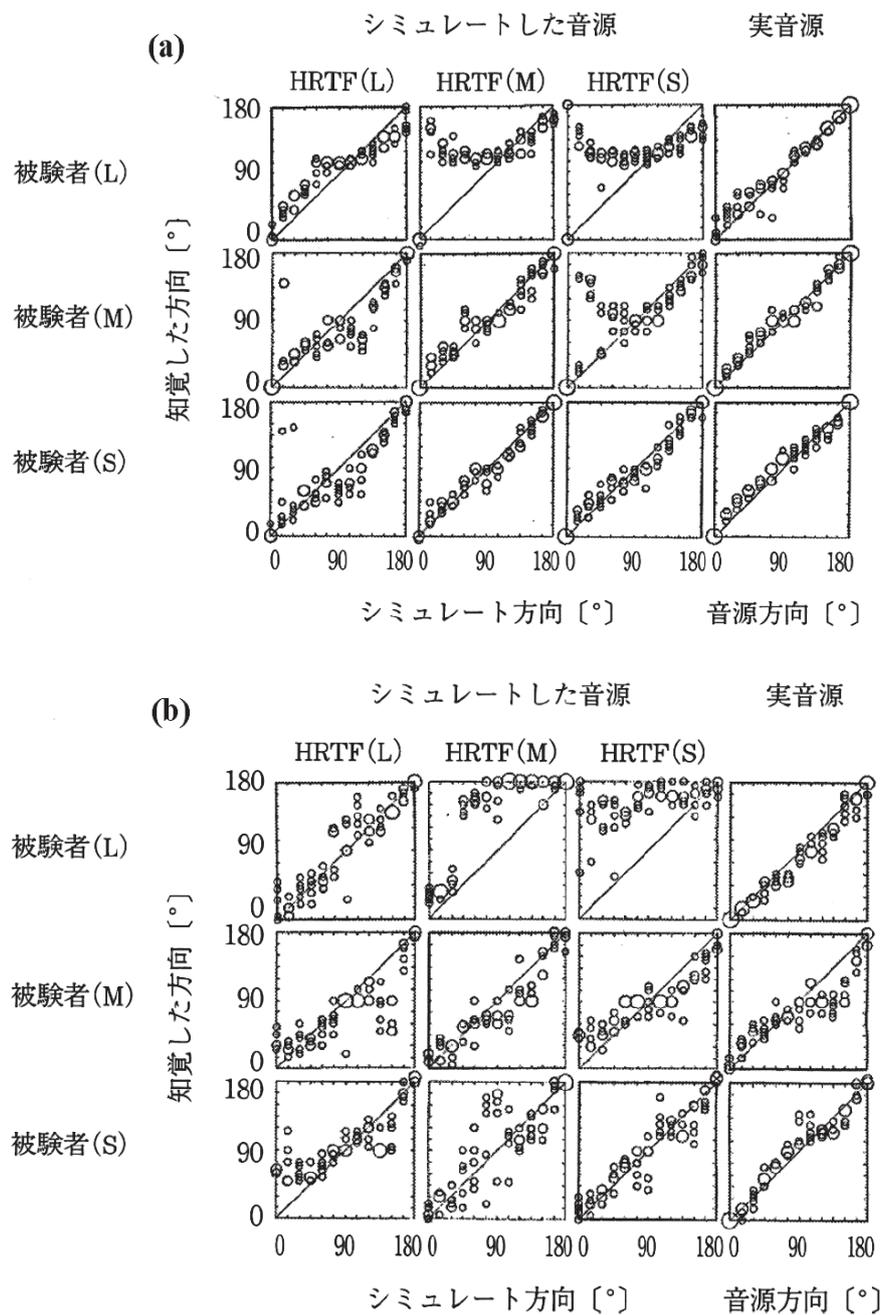


図 1.12 本人および他人の HRTF を用いた水平面(a), 正中面(b)での
音像定位実験の結果[2]

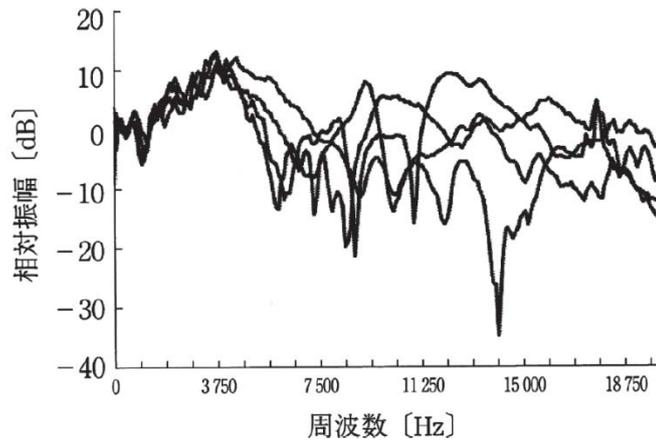


図 1.13 正面方向の HRTF [7]

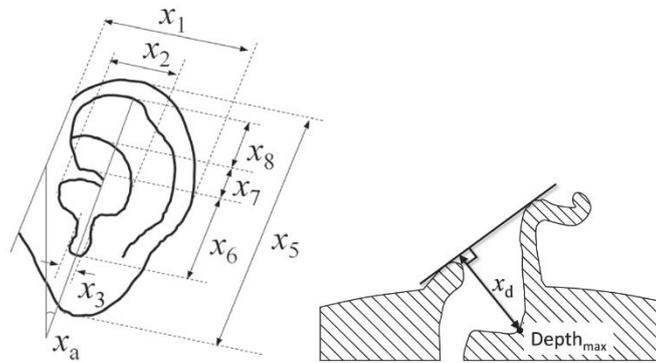


図 1.14 耳介形状 [45]

以上より，精度の高い 3 次元音像制御を実現するには，受聴者本人の HRTF，あるいは適合する HRTF が必要となる．受聴者本人の HRTF は測定や数値計算で得ることが可能であるが，無響室や MRI などの特殊な設備が必要であり，全ての受聴者の HRTF を求めることは実用化を考えると現実的な方法とは言えない．精度の高い 3 次元音像制御が一般に普及するためには，受聴者に適合する HRTF を容易に求める方法が必要である．

1.5 受聴者に適合する頭部伝達関数の提供方法

従来提案されている，受聴者に適合する HRTF の提供方法を図 1.15 にまとめた．大きな方針の違いとして，HRTF を標準化する考え方と，HRTF を個人化する考え方がある．1.5.1 項では HRTF の標準化に関する従来の取り組みを，1.5.2 項では HRTF の個人

化に関する従来の取り組みを概説する。なお、本論文で提案する HRTF の個人化方法は、HRTF の個人化(1.5.2 項)方法のうち、方向知覚の手掛かりを適合させる方法(1.5.2(B-4))とデータベースから適合する HRTF を選出する方法(1.5.2(B-3))を組み合わせたものである。

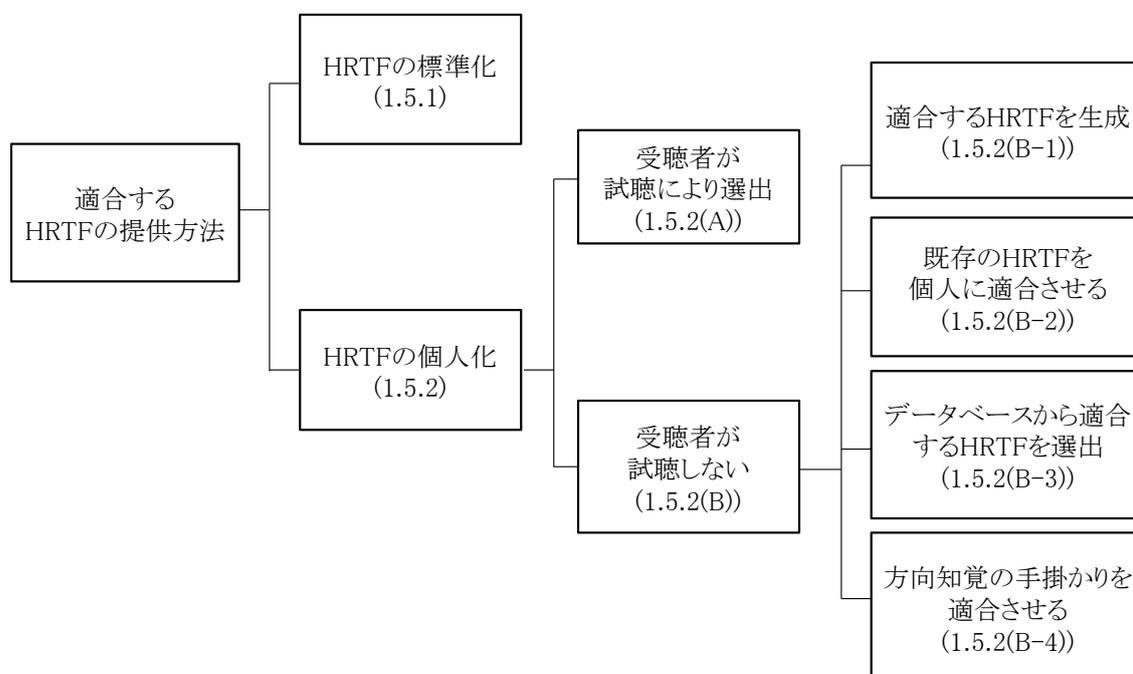


図 1.15 受聴者に適合する HRTF の提供方法.

図中の () 内は本論文中の項を表す.

1.5.1 頭部伝達関数の標準化

HRTF を標準化する方式については、ハードウェアとして実現したダミーヘッドがある。KEMAR(Knowles electric manikin for acoustic research)は、多人数の頭部、耳介、胴体の寸法を計測し、その中央値を用いて作成された[52]。しかし、多くの受聴者では音像の前後誤判定や上昇、頭内定位などが発生する。これは、N1,N2 周波数の弁別閾(0.1-0.2 oct.[51])に対し、その個人差は 1.27, 1.17 oct.に及ぶこと[45]から説明できる。つまり、標準化した HRTF は、HRTF の個人差を解決しているとはいえない。

1.5.2 頭部伝達関数の個人化

HRTF を個人化 (Personalization もしくは Individualization) する方法については、受聴者が試聴により適合する HRTF を選出する方法、および受聴者の試聴を必要としない方

法に大別できる。

A) 受聴者が試聴により適合する HRTF を選出する方法

この方法では、データベースの規模が大きくなるほど試聴に要する時間が増えるため、時間を短縮する方法が検討されている。Seeber and Fastl [53]は、2ステップ選択法を提案している。この方法の第1ステップでは、HRTFのグループを選定し、第2ステップでそのグループの中から最良の1つを選択する。この方法では約10分で受聴者に適合するHRTFを選択できると報告している。Iwaya [54]はトーナメント方式による試聴方法を提案している。受聴者は、HRTFを用いて作成された仮想音源を試聴して最も適合するHRTFを選出する。32名分のHRTFから受聴者に適合するHRTFを選出するのに要する時間は、約15分であると報告している。しかし、これらの方法で用いるHRTFは、データベースから無作為に選出されたものであり、受聴者に適合するHRTFが含まれている保証がない。

B) 受聴者の試聴を必要としない方法

この方法は、さらに以下の方針に大別できる。

- 1) 個人に適合する HRTF を生成する
- 2) 既存の HRTF を個人に適合させる
- 3) データベースから適合する HRTF を選出する
- 4) 方向知覚の手掛かりを個人に適合させる方法

B-1) 個人に適合する HRTF を生成する方法

Kistler and Wightman[55]は、PCA(principal component analysis)によって、HRTFをいくつかの主成分に分解できることを示し、最初の5個の主成分でHRTFの振幅スペクトルの90%が再現できるとした。PCAを利用して、受聴者の耳介形状から各主成分の係数を推定し、個人化HRTFを合成する試みがある[56]。しかし、いずれの文献においても、この方法により個人化したHRTFの音像定位精度は示されておらず、その有効性は明らかではない。また、寄与率の高い主成分がスペクトラルキューに対応しているとは限らないため、スペクトラルキューとして重要なノッチを再現するには不十分であり、ノッチを再現するためには主成分が15-20程度必要であるとする計算例もある。

B-2) 既存の HRTF を個人に適合させる方法

Middlebrooks[57-59]は、標準的な HRTF を受聴者の耳介形状に応じて周波数軸上で伸縮する方法を提案した。彼は、代表的な HRTF の方向依存性(DTF: directional transfer function)を周波数軸上で適切に伸縮することで、受聴者本人の DTF に近似することが可能であり、音像定位実験により、他人の DTF を適切に伸縮した DTF の quadrant error (前後もしくは上下方向における 90°以上の誤差のことで、前後誤判定率、上下誤判定率を足し合わせたようなもの) が 14.7%であり、被験者本人の DTF の quadrant error 15.6%と同等であることを示した。ただし、受聴者に適合した伸縮率を見いだすには、1ブロック 20分を要する試聴実験を1から3ブロック実施する必要があるという問題があり、適切な伸縮率を受聴者の耳介形状から導くことには成功していない。

B-3) データベースから適合する HRTF を選出する方法

この方法の代表的な取り組みとして、Zotkin *et al.* [60]が提案している。このアプローチは、耳介の形状が似ていれば、HRTF も似ているという仮説に支えられている。彼らは、8人の受聴者について耳介部位7箇所を計測し、その計測値に最も近い耳介の HRTF をデータベースから選出した。この HRTF を用いて前半球面を目標方向とした音像定位実験を行ったが、その音像定位精度はダミーヘッドの HRTF と比較してわずかに 1.9°改善されただけであった。この方法の問題点として、すべての耳介部位の個人差を同じ重みで扱っていることが挙げられる。HRTF のピークは耳甲介腔の深さ方向のモード、耳介の上下方向のモードに由来し、ノッチは外耳道入口付近に生じた節に由来することから、これらの現象に関連する部位に着目して耳介形状の類似度を評価する必要がある。

B-4) 方向知覚の手掛かりを個人に適合させる方法

水平面内の両耳間差キューを推定する方法について提案されている。

受聴者の頭部形状から ITD を推定する方法については、Algazi *et al.* [61]および渡邊ら [62]が提案している。Algazi *et al.* [61]は、頭部を完全な球体と仮定して受聴者の頭部形状から適切な球の半径を推定し、その球の半径から各方向の ITD を算出する方法を提案した。彼らは、推定した球の半径と受聴者本人に適した球の半径の誤差が平均で 1.2 mm であったと報告している。また、被験者本人の方位角に対する ITD の振舞いを数式でモデル化した結果、モデル化誤差は側方において約 12°であったと報告している。方

向知覚の弁別閾が側方で約 10° [63]であることから、方位角に対する ITD の振舞いを十分な精度でモデル化できているとはいえない。渡邊ら[62]は、方位角に対する ITD の振舞いを正弦関数の和でモデル化し、その正弦関数のパラメータを受聴者の頭部形状から推定する方法を提案した。ただし、この正弦関数の和によって ITD をモデル化する際に $25 \mu\text{s}$ 程度の誤差が生じる。

ILD については、Watanabe *et al.* [64]は 1/3 octave band ごとの水平面内の ILD を正弦関数の和でモデル化し、その係数を頭部形状から推定する方法を提案し、4名の被験者を用いて ILD の推定精度を検証した。水平面内 72 方向の平均推定誤差を被験者ごとに求めた結果、4名の被験者の最大推定誤差は 1,250 Hz 以下の帯域では約 2–3 dB、それ以上の帯域では 3–8 dB であり、全ての帯域で方位角に関する ILD の弁別閾(約 1 dB[65])を超えていた。

1.6 3次元音像制御に必要な頭部伝達関数の情報

個人に適合する HRTF を再現することでその方向に音像を制御することができる。言い換えれば、目標とする全ての方向について、適合する HRTF を求める必要がある。この方向数を削減するために、以下の知見が有用であると考えられる。

従来、3次元音像制御に必要となる HRTF の個数を削減することを目的として、正中面内の実測 HRTF と水平面内の両耳間差の組み合わせる方法が提案されている[66,67]。Morimoto and Aokata[66]は、音源の方向を側方角 α と上昇角 β によって表わした矢状面座標系(図 1.1)を提案した。側方角が等しい点を結ぶ円周(矢状面)上では、両耳間時間差(ITD)および両耳間レベル差(ILD)はほぼ一定となる。さらに、彼らは側方角 α と上昇角 β をそれぞれ両耳間差キューとスペクトラルキューによって独立に制御できるのではないかと提唱した。この考えに基づき、Morimoto *et al.* [67]は、被験者本人の上半球正中面内の実測 HRTF と水平面内の両耳間差を用いて、側方角 $\alpha=0-90, 30^\circ$ 間隔の矢状面において、上昇角 $\beta=0-180^\circ$ (30° 間隔)を目標方向とした音像定位実験を行った。その結果、半球面内の任意の方向に音像制御が可能である。すなわち、側方角と上昇角は、それぞれ水平面内の両耳間差と正中面内の実測 HRTF によって独立に制御できることを示した。

また、P1,N1,N2 で再構成した正中面内のパラメトリック HRTF と水平面内の両耳間差キューを組み合わせることで、一部の被験者の上方を除き、被験者本人の HRTF と同

等の定位精度となることが示されている。

これらの結果は、HRTF の個人差を以下の情報に限定できる可能性を示唆している。

- 1) 正中面内のスペクトラルキュー
- 2) 水平面内の両耳間差キュー

1.7 本論文で取り扱う研究課題

耳入力信号再現方式の普及の阻害要因の 1 つは HRTF の個人差にあると考えられる。したがって、HRTF の個人化方法を確立する必要がある。ただし、HRTF の個人化方法を一般に普及させるためには、以下のように簡易的な方法である必要があると考えられる。

- ・特別な訓練や設備を必要としない。
- ・多大な時間を要しない。
- ・受聴者が試聴する必要がない。

また、N1, N2 がスペクトラルキューとして重要な役割を果たすが、N1,N2,P1 を再現した場合、一部の被験者において上方で音像定位精度が低下する場合があった。したがって、上方における N1,N2 以外のスペクトラルキューの役割を解明することで、HRTF の個人化において推定すべきスペクトラルキューを特定できる可能性がある。

1.8 本研究の目的

本研究では、高い定位精度が得られる 3 次元音像制御を一般に普及させるため、以下の 2 つの研究目的を有する。

1 つめは、HRTF の個人化方法の確立である。本研究が目指す HRTF の個人化方法のプロセスを図 1.16 に示す。提案法は、方向知覚の手掛かり、およびその成因について着目した方法である。まず、頭部形状から水平面内の ITD および ILD をそれぞれ推定することで、両耳間差キューの個人化を目指す。さらに、耳介形状から N1,N2 周波数を推定し、それに最も近い HRTF をデータベースから選出することでスペクトラルキューの個人化を目指す。これらを組み合わせることで精度の高い 3 次元音像制御を提供できると考えられる。本研究では、HRTF の個人化精度の目標値は、本人 HRTF の音像定位精度とする。すなわち、ITD および ILD については、方向知覚の弁別閾内の推定誤差を目標値とする。また、スペクトラルキューについては、本人 HRTF と同等の上昇角定位精

度を目標値とし，データベースから選出した HRTF を用いて音像定位実験を行う．

2つめは，上方における音像定位精度の向上である．Iida *et al.*[15]は，N1,N2 を再現した音像定位実験の結果，一部の被験者において上方で定位精度が落ちる場合があることを報告している．そこで，本研究では，上方におけるスペクトラルキューに関する検証を行い，上方における音像定位精度の向上を目指した．

図 1.17 に本論文の章構成を示す．第 2 章，第 3 章では，両耳間差キューとスペクトラルキューの個人化方法を提案し，その有効性を示す．第 4 章では，上方におけるスペクトラルキューに関する検証を行う．

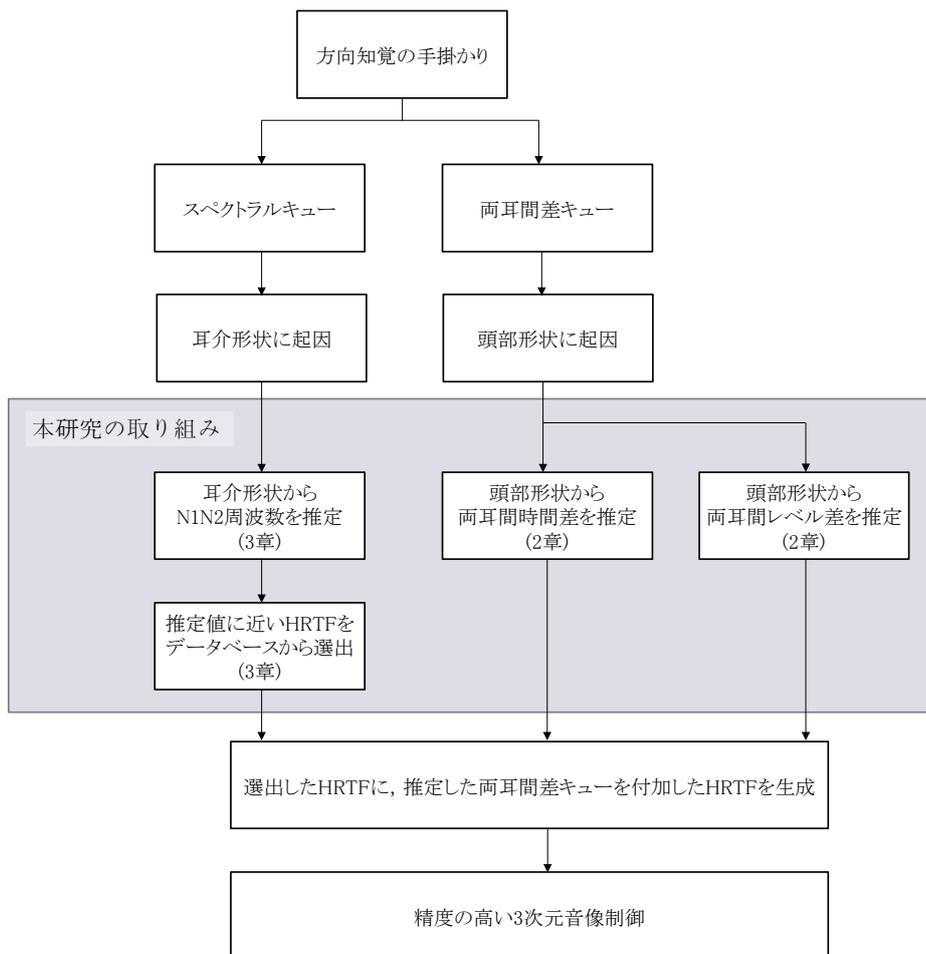


図 1.16 本研究における HRTF の個人化方法のプロセス

図中の () 内は本論文中の章を表す。

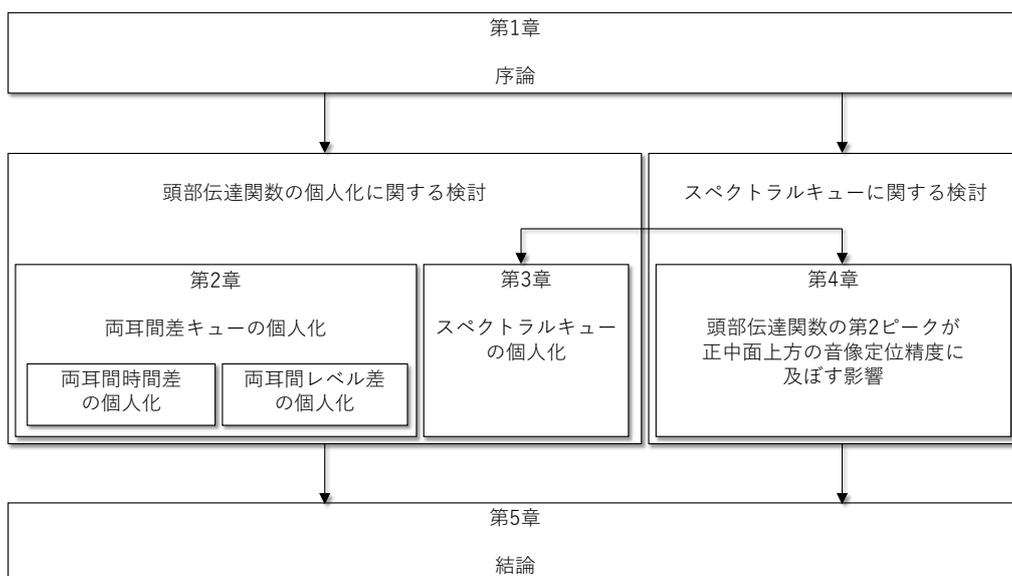


図 1.17 本論文の章構成

第2章 両耳間差キューの個人化

2.1 本章の目的

HRTFには個人差が存在するため、他人のHRTFを用いると誤った方向に定位する現象がしばしば発生する。本章では、左右方向知覚の手掛かりであるITDとILDの個人差について採り上げる。受聴者の頭部形状からITDとILDをそれぞれ弁別閾内の精度で推定できれば、高い精度で左右方向の音像制御が実現できる。

従来、ITDとILDをそれぞれ受聴者の頭部形状から推定する方法が提案されている[61,62,64]。しかし、ITDについては、Algazi *et al.*[61]の方法では、側方において方向知覚の弁別閾を超えていた。また、渡邊ら[62]の方法では、ITDの振舞いをモデル化する際の誤差が25 μ s程度あることが報告されている。一方、ILDについては、Watanabe *et al.*[64]の方法では、全ての帯域のILDの推定誤差が弁別閾を超えていた。すなわち、従来法では、受聴者の頭部形状に基づいて左右方向知覚の手掛かりであるITDとILDを弁別閾内の精度で推定できていない。

本章では、各受聴者の両耳間差キューを弁別閾内の精度で推定することを目指し、受聴者の頭部形状から両耳間差キューを推定する方法を提案した。

2.2 両耳間時間差の個人化

2.2.1 従来の研究

受聴者の頭部形状からITDを推定する方法については、Algazi *et al.* [61]および渡邊ら[62]が提案している。Algazi *et al.* [61]は、頭部を完全な球体と仮定して受聴者の頭部形状から適切な球の半径を推定し、その球の半径から各方向のITDを算出する方法を提案した。彼らは、推定した球の半径と受聴者本人に適した球の半径の誤差が平均で1.2 mmであったと報告している。また、被験者本人の方位角に対するITDの振舞いを数式でモデル化した結果、モデル化誤差は側方において約12°であったと報告している。方向知覚の弁別閾が側方で約10°であることから、方位角に対するITDの振舞いを十分な精度でモデル化できているとはいえない。渡邊ら[62]は、方位角に対するITDの振舞いを正弦関数の和でモデル化し、その正弦関数のパラメータを受聴者の頭部形状から推定する方法を提案した。ただし、この正弦関数の和によってITDをモデル化する際に25 μ s程度の誤差が生じる。彼らが用いた頭部モデルは、Algazi *et al.* [61]と同様に頭部の左

右非対称性を考慮していないため ITD の左右非対称性に対応できていない。

従来法には以下の解決すべき課題がある。

- 1) 頭部形状の前後・左右非対称性が考慮されていないため、ITD の前後・左右非対称性に対応できない。
- 2) 各頭部形状の回帰係数を方向間で等しく扱っているため、方向ごとの ITD と各頭部形状の相関関係の違いに対応できない。

2.2.2 本節の目的

本節では、頭部の前後・左右非対称性に対応するモデルを導入することで水平面内の ITD の推定精度の向上を図った。また、目的変数を水平面内の ITD、説明変数を頭部形状として方向ごとに異なる重回帰式を構築することで方向ごとの ITD と頭部形状の相関関係の違いに対応させた。

2.2.3 両耳間時間差の算出

a) 頭部インパルス応答の測定

ITD は、無響室において測定した被験者本人の hrir(head-related impulse response)から算出した。hrir の測定方向は水平面内 12 方向 (30°間隔)、スピーカから頭部中心位置までの距離は 1.2 m である。スピーカから頭部中心位置までの距離が 1 m を超える場合、HRTF には距離依存性が生じないことが示されている[2]。したがって、従来法(1.5 m)と提案法(1.2 m)の hrir の測定距離の差は無視できると考えられる。

実験システムは、ノート PC、オーディオインタフェース(RME FireFaceUFX)、スピーカ(FOSTEX FE83E)、アンプ(Marantz PM4001)、A/D コンバータ(Roland EDIROL M-10MX)、耳栓型マイクロホン[15]で構成した(図 2.1)。測定信号は Swept-Sine Signal (2¹⁸ サンプル)である。測定信号および録音のサンプリング周波数は 48 kHz である。測定時、被験者に対して椅子に座り背筋を伸ばして正面を向いて静止するように指示した。被験者の頭部は固定していないが頭部および胴体に傾きがないかについて、被験者の前方および側方から 2 名の実験者が目測で確認した。被験者は 20 代の 33 名(男性 27 名、女性 6 名)である。

hrir の算出方法は以下の通りである。

- 1) スピーカから受聴者の左右の外耳道入口(外耳道閉塞状態[16])までのインパルス

- 応答 $g_l(t)$, $g_r(t)$ を測定する (2^{18} サンプル).
- 2) 受聴者がいない状態での音源から受聴者の頭部中心位置に相当する位置までのインパルス応答 $f(t)$ を測定する (2^{18} サンプル).
 - 3) 正面方向の $g_l(t)$ と $g_r(t)$ から振幅の絶対値が最大となる時刻をそれぞれ求め, 時刻が早い方から 50 サンプル前を切り出し開始位置 T_0 とする.
 - 4) 全方向の $g_{l,r}(t)$ および $f(t)$ を T_0 から 512 サンプルの矩形窓で切り出す.
 - 5) $g_{l,r}(t)$ および $f(t)$ をフーリエ変換し, 512 サンプルの複素数 $G_{l,r}(\omega)$, $F(\omega)$ を得る.
 - 6) 式(1.1)により $HRTF(\omega)$ を求める.
 - 7) $HRTF(\omega)$ を逆フーリエ変換し, $hrir(t)$ を得る.

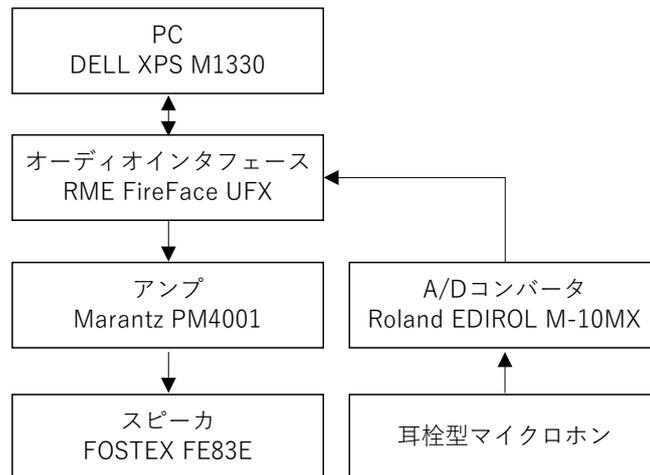


図 2.1 hrir の測定機器の系統図

b) 両耳間時間差の算出方法

2.2.3(a)の方法で求めた $hrir$ を用いて ITD を算出した. ITD の算出手順は以下の通りである.

- 1) $hrir$ に 1.6 kHz をカットオフ周波数とする最小位相系ローパスフィルタ(790 tap)を畳み込む.
- 2) サンプリング周波数を 8 倍($48 \text{ kHz} \times 8 = 384 \text{ kHz}$)に変換する (時間分解能 $2.604 \mu\text{s}$).
- 3) 両耳間相互相関関数 Φ (式 2.1) が最大となる時間差 τ を ITD とする[68].

$$\Phi_{l,r}(\tau) = \lim_{T \rightarrow \infty} \frac{\int_{-T}^T hrir_l(t) \times hrir_r(t - \tau) dt}{\sqrt{\int_{-T}^T hrir_l^2(t) \times hrir_r^2(t) dt}} \quad (2.1)$$

ここで、 $|\tau| \leq 1000 \mu\text{s}$ とし、添え字の l, r はそれぞれ左耳、右耳を表す。

c) 両耳間時間差の算出結果

33名の ITD を図 2.2 に示す。音源の方位角は正面方向を 0° として時計回りに表した。正の値は左耳の到達時刻が遅いことを表す。 0° (正面) および 180° (真後ろ) において $0 \mu\text{s}$ とならない場合があるが、これは頭部の左右非対称性が影響していると考えられる。いずれの被験者においても、ITD は真横($90, 270^\circ$)で最大となった。真横から $\pm 30^\circ$ において、個人差が大きい傾向にある。

表 2.1 に各方向における ITD の平均値(Ave.)、最大値(Max)、最小値(Min)、個人差の範囲(Max-Min)、標準偏差(Std.)を示す。Ave.をみると、真横($90, 270^\circ$)では約 $730 \mu\text{s}$ となった。同一矢状面内 (30° と 150° , 60° と 120° , 330° と 210° , 300° と 240°) で比較すると、前方の方が $15.1-47.1 \mu\text{s}$ 大きい。これは、外耳道入口が頭部中心よりやや後方に位置している可能性が示唆される。

次に、ITD の分布範囲(Max, Min)をみる。 90° の ITD の分布範囲は 60° のそれと互いに重なっていた。また、 270° の ITD の分布範囲は $240, 300^\circ$ のそれと互いに重なり合っていた。つまり、ITD には 30° 分以上の個人差が存在する。方向知覚の弁別閾が前方では 1° 、側方では 10° 程度[63]であることから、ITD には方向知覚に影響を及ぼす個人差が存在すると言える。

Max-Min は 12 方向において $83.3-148.4 \mu\text{s}$ の範囲をとり、真横 $\pm 30^\circ$ ($60, 120, 240, 300^\circ$)で大きく、真横($90, 270^\circ$)で小さい傾向にある。全方向の平均値は $107.9 \mu\text{s}$ であった。Std.は $19.6-35.2 \mu\text{s}$ の範囲をとり、Max-Min と同様に真横 $\pm 30^\circ$ ($60, 120, 240, 300^\circ$)で大きく、真横($90, 270^\circ$)で小さい傾向にある。以上より、ITD の個人差は、真横 $\pm 30^\circ$ ($60, 120, 240, 300^\circ$)で大きいと言える。

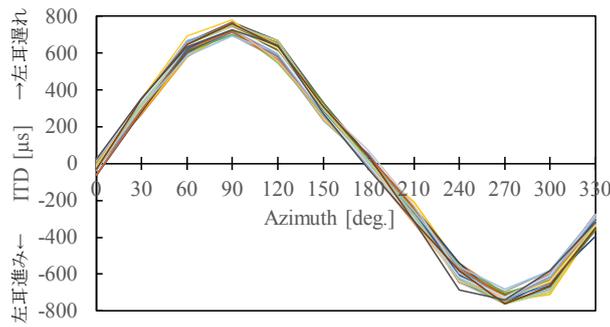


図 2.2 全被験者の ITD の方位角に対する変化

表 2.1 33 名の ITD の統計値 [μs]

Azm. [deg.]	Ave.	Max	Min	Max-Min	Std.
0	-14.7	26.0	-65.1	91.1	22.5
30	310.8	354.2	265.6	88.5	24.2
60	629.2	692.7	578.1	114.6	25.7
90	733.2	778.6	692.7	85.9	19.6
120	614.1	669.3	546.9	122.4	35.2
150	280.7	330.7	234.4	96.4	22.5
180	13.3	70.3	-28.6	99.0	23.1
210	-270.4	-210.9	-325.5	114.6	28.8
240	-602.1	-539.1	-687.5	148.4	35.0
270	-734.7	-679.7	-763.0	83.3	21.1
300	-638.1	-580.7	-710.9	130.2	34.2
330	-317.5	-273.4	-393.2	119.8	27.8
Ave.	-	-	-	107.9	26.6

2.2.4 頭部寸法の特徴量の計測

a) 頭部形状の計測方法

頭部表面を通る音波の伝播経路の個人差が ITD の個人差に対応すると考え、ITD を求めた 33 名の頭部形状を計測した。

従来、Algazi *et al.* [61]は頭の幅、奥行、高さの 3 箇所を計測しており、頭の幅と奥行が ITD の推定への寄与率が高いことを示している。渡邊ら[62]は、Algazi *et al.* [61]が用いた 3 箇所に加え、鼻の高さ、長さ、奥行、耳介の幅、長さ、肩幅、頭耳長、頭耳高を加えた 11 箇所を用いている。このうち、頭の幅、高さ、奥行、頭耳長、頭耳高、肩幅が ITD の推定への寄与率が高いことを示している。ただし、いずれの研究においても、頭部の前後・左右非対称性は考慮されていない。

本研究で用いる頭部モデルを図 2.3、表 2.2 に示す。従来法で ITD の推定への寄与率が高いことが示された部位（頭の幅、高さ、奥行、頭耳長、頭耳高、肩幅）は全て用い

ている。このうち、 $p4_{l,r}$, $p5_{l,r}$ は渡邊ら[62]の頭耳長に相当する部位であるが、頭部の前後・左右非対称性に対応できるように前後・左右に分割した。また、 $p6_{l,r}$ は渡邊ら[62]の頭耳高に相当する部位であるが左右に分割した。なお、 $p4_{l,r}$, $p5_{l,r}$ の始点と終点は外耳道入口の中心位置、および正中面と水平面の交点であり、 $p6_{l,r}$ の始点と終点は外耳道入口の中心位置、および正中面と横断面の交点である。

頭部形状の計測には触角計と巻尺を用いた。p1, p2, p3, p7 は触角計、p4, p5, p6 は巻尺により計測した。最小読取値は触覚計、巻尺ともに 1 mm である。

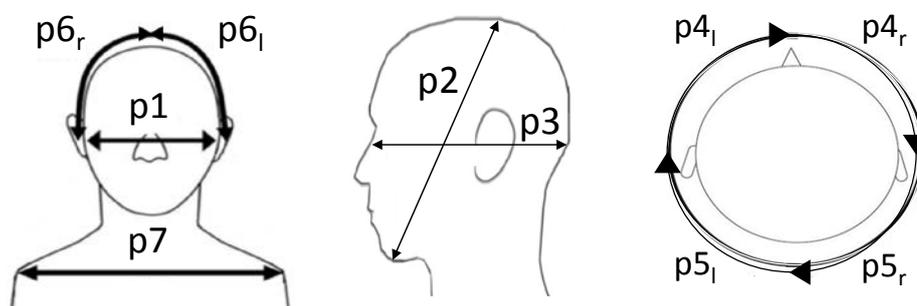


図 2.3 ITD の推定に用いた頭部寸法の計測箇所

表 2.2 計測箇所とその名称

計測箇所	名称
p1	頭の幅
p2	頭の高さ
p3	頭の奥行
p4, p5	頭周
p6	頭頂から外耳道入口
p7	肩幅

b) 頭部形状の計測結果

表 2.3 に頭部形状の計測結果と従来法[61,62]の計測結果の平均値を併せて示す。本研究で計測した p1, p3, p7 は従来法の平均値に比べやや小さく、p2 はやや大きい。

個人差の範囲(Max-Min)は、p7 (肩幅) が最も大きく 102 mm であった。一方、p1 (頭の幅) が最も小さく 18 mm であった。その他の部位は 26–56 mm の個人差が存在した。

表の RSD(Relative Standard Deviation)は標準偏差(Std.)を平均値(Ave.)で割ったものであ

り，相対的なばらつきを表す指標である．頭部形状の個人差を RSD で評価すると， $p4_{l,r}$ ， $p5_{l,r}$ （頭周）と $p7$ （肩幅）の個人差が大きい．

次に，頭部の前後・左右非対称性についてみる． $p4_{l,r}$ ， $p5_{l,r}$ （頭周）の平均値の前後差については， $p4_l$ ， $p5_l$ （左側）では平均で $p4_l$ （前方）が $p5_l$ （後方）より 8 mm 大きく， $p4_r$ ， $p5_r$ （右側）も $p4_r$ （前方）が $p5_r$ （後方）より 14 mm 大きい．この結果から，外耳道入口は頭部中心よりやや後方に位置していると考えられ，左右の耳への音波の経路差が後方よりも前方の方が大きくなると推測できる．これは，2.3 節で示した同矢状面内の ITD が後方よりも前方の方が大きくなったことと合致する．左右差についてみると， $p4_{l,r}$ ， $p5_{l,r}$ ， $p6_{l,r}$ の平均値は左側の方がそれぞれ 7, 1, 2 mm 小さい．

表 2.3 33 名の頭部形状の統計値

	Ave.	Max	Min	Max-Min	Std	RSD	渡邊ら (Ave.)	Algazi <i>et al.</i> (Ave.)	
								Women	Men
p1	143	152	134	18	5.6	0.039	156	145	155
p2	246	266	227	39	9.9	0.040	227	-	-
p3	185	201	170	31	7.9	0.043	199	180	196
$p4_l$	152	167	136	31	7.4	0.049	-	-	-
$p4_r$	159	185	148	37	8.1	0.051	-	-	-
$p5_l$	144	163	130	33	8.1	0.056	-	-	-
$p5_r$	145	176	120	56	12.0	0.083	-	-	-
$p6_l$	195	217	176	41	8.2	0.042	-	-	-
$p6_r$	197	211	185	26	7.2	0.036	-	-	-
p7	393	439	337	102	23.9	0.061	408	-	-

2.2.5 単回帰分析

音源方向ごとに頭部形状と ITD の単回帰分析を行った．頭部形状と各方向における ITD の単相関係数を表 2.4 に示す．表中の網掛けは有意水準 10% で相関関係が認められることを表す($0.29 \leq |r|$)．ITD と相関があった主な部位は以下の通りである．

$p1$ （頭の幅）は，最も多くの方向（12 方向中 6 方向）で相関関係が認められ， $90, 270^\circ$ において強い相関関係($0.6 < |r|$)となった． $p1$ は従来法[61,62]で ITD の推定への寄与率が最も高いと報告された部位であり，本研究においても同様の結果となった．

$p2$ （頭の高さ）および $p3$ （頭の奥行）は，側方で有意な相関関係となる傾向がある．

$p4_{l,r}$ ， $p5_{l,r}$ （頭周）については， $p4_{l,r}$ は音源が水平面内の前半面にある場合のみ有意な相関が認められ， $p5_{l,r}$ は音源が水平面内の後半面にある場合のみ有意な相関が認められた．

つまり、音源から近い側の頭周の個人差は遠い側の頭周の個人差に比べ、ITDの個人差との相関が強いと言える。

p6_{l,r} (頭頂から外耳道入口) については、p6_lは150°でのみ相関関係が認められた。p6_rは60, 90, 270, 300°で相関が認められ、p6_rの90, 270°で強い相関関係(0.6 < |r|)となった。

p7 (肩幅) は270, 300°でのみ相関関係が認められた。

方位角ごとにみると、側方において相関関係が認められる部位が多い傾向にあり、真横(90, 270°)においてはp1 (頭の幅) の相関が最も高かった(r=0.61, -0.65)。しかし、30°と210°では、相関関係が認められる部位は存在しなかった。したがって、単回帰分析では全ての方向のITDを精度良く推定することは困難である。

表 2.4 頭部形状と各方向のITDの単相関係数 r. : 10%有意(0.29 ≤ |r|)

Azm. [deg.]	p1	p2	p3	p4 _l	p4 _r	p5 _l	p5 _r	p6 _l	p6 _r	p7
0	-0.45	-0.20	-0.06	-0.40	-0.28	0.19	-0.09	-0.18	-0.17	-0.14
30	0.20	0.05	0.10	0.02	0.18	-0.04	0.01	-0.04	0.16	0.03
60	0.54	0.34	0.41	0.29	0.44	0.11	0.12	0.20	0.39	0.28
90	0.61	0.34	0.57	0.43	0.50	0.16	0.37	0.27	0.61	0.24
120	0.10	-0.03	0.38	0.04	0.02	0.28	0.38	0.09	0.22	-0.06
150	0.01	-0.02	0.35	-0.10	-0.05	0.47	0.21	-0.29	0.14	-0.02
180	0.35	0.24	0.25	0.21	0.26	-0.17	0.11	0.14	0.14	0.05
210	0.00	0.20	0.05	0.12	0.07	-0.03	-0.02	0.25	0.06	-0.01
240	-0.23	-0.13	-0.23	-0.07	-0.16	-0.32	-0.09	-0.20	-0.26	-0.10
270	-0.65	-0.33	-0.46	-0.52	-0.58	-0.21	-0.34	-0.07	-0.64	-0.46
300	-0.54	-0.38	-0.42	-0.42	-0.63	0.15	-0.18	-0.01	-0.41	-0.51
330	-0.25	-0.16	-0.16	-0.19	-0.31	-0.09	0.02	0.08	-0.08	-0.02

2.2.6 重回帰分析

10箇所(10変数)の頭部形状を説明変数、各方向のITDを目的変数とした重回帰分析を行った(式2.2)。

$$ITD(s, \phi) = a_1(\phi)p_1 + a_2(\phi)p_2 + \dots + a_7(\phi)p_7 + \beta(\phi) \quad [\mu s] \quad (2.2)$$

ここで、sは被験者、φは方位角である。

多重共線性の有無については、VIF(Variance Inflation Factor)によって確認した。VIFは以下の式で定義される。

$$VIF(j) = \frac{1}{(1-R^2(j))} \quad (2.3)$$

ここで、R²(j)はj番目の説明変数を目的変数とし、他の説明変数を説明変数とした場

合の重回帰分析における決定係数である。VIF（10 種類）の値が 10 未満であれば説明変数間には多重共線性はないとみなせるが[69]，本研究で用いたすべての説明変数の VIF は 10 未満であった。

表 2.5 に重相関係数 r ，回帰モデル全体の危険率 p ，残差の絶対値の平均 E と回帰係数（ β は定数項）を示す。各方向における重相関係数 r は 0.34–0.79 であり，側方で高くなる傾向がある。全ての方向において有意水準 5%で相関関係が認められた($0.34 \leq r$)。全方向の平均重相関係数は 0.59 であった。

危険率 p については，90, 150, 270, 300°で 10%未満となった。ただし，単回帰分析で相関関係が認められる部位が存在しなかった 30, 210°では高い危険率(97.5, 94.9%)となった。

各方向における残差の絶対値の平均 E は 9.6–24.8 μs の範囲をとった。全方向の平均は 16.5 μs であり，表 2.1 に示す ITD の個人差の範囲の全方向の平均 107.9 μs の 15.3%に相当する。残差は真横(90, 270°)では小さく(9.8, 9.6 μs)，120, 240°では大きい傾向(20.8, 24.8 μs)にあるが，全ての方向において渡邊ら[62]の方法による ITD のモデル化誤差(25 μs)よりも小さい。したがって，本研究の重回帰モデルによって，従来法よりも高い精度で受聴者の ITD を推定できる可能性が示唆される。

表 2.5 重回帰分析の結果. r: 重相関係数, p: 危険率, E: 残差の絶対値の平均, β : 定数項 (頭部形状の単位は mm)

Azm. [deg.]	r	p	E [μ s]	Regression coefficient ($\times 10^{-3}$)										
				α_1	α_2	α_3	α_{4_l}	α_{4_r}	α_{5_l}	α_{5_r}	α_{6_l}	α_{6_r}	α_7	β
0	0.59	0.363	14.2	2.34	0.06	-0.72	1.11	-0.56	-0.48	-0.13	-0.21	-0.15	0.01	-127.24
30	0.34	0.975	17.9	-1.37	-0.34	-0.19	0.98	-0.33	0.24	0.18	0.43	-0.51	0.16	-197.13
60	0.63	0.210	15.0	-2.60	-0.27	-1.05	1.16	-0.02	0.20	0.55	-0.02	-0.62	0.02	-160.29
90	0.78	0.009	9.8	-0.92	0.08	-1.13	0.40	-0.30	0.21	-0.04	-0.18	-1.13	0.22	-278.76
120	0.65	0.165	20.8	1.50	1.23	-3.11	0.32	0.11	-0.15	-0.67	-0.83	-1.01	0.40	-295.37
150	0.71	0.057	12.2	-0.09	-0.27	-1.32	0.48	0.07	-1.09	-0.03	1.10	-0.41	0.39	-163.68
180	0.50	0.672	16.0	-1.81	-0.61	-0.82	0.63	0.15	0.79	0.11	0.06	-0.03	0.28	185.10
210	0.38	0.949	21.1	1.59	-1.18	0.53	0.15	-0.90	-0.04	-0.41	-0.71	-0.15	0.33	458.68
240	0.43	0.874	24.8	0.79	0.15	0.63	-1.18	0.41	0.98	-0.17	0.53	0.71	-0.10	127.05
270	0.79	0.005	9.6	1.18	-0.14	-0.33	0.39	0.66	0.35	0.10	-0.59	1.07	0.04	320.99
300	0.75	0.022	18.4	2.00	-0.26	0.67	-0.72	1.08	-1.24	-0.37	-0.55	0.91	0.42	217.87
330	0.54	0.529	17.6	1.00	1.70	-1.24	-0.09	1.84	1.06	0.22	-1.46	-0.38	-0.67	142.87
Ave.	0.59	0.402	16.5											

2.2.7 ナイーブな被験者による両耳間時間差の推定精度の検証

前項で求めた重回帰モデル（表 2.5）の推定精度を重回帰分析に含めていない被験者を用いて検証した。被験者は 20 代の女性 2 名(A, B)と男性 2 名(C, D)の 4 名である。この 4 名の ITD と頭部形状を求めた（表 2.6, 2.7）。

重回帰式を用いて頭部形状から 12 方向の ITD を推定した。表 2.8 にナイーブな被験者の推定誤差の絶対値（以降、推定誤差）を示す。以降、各方向の重回帰モデルの残差の絶対値の平均（以降、残差）を基準として推定精度を評価する。被験者 A の推定誤差は 90–150°を除き、各方向の残差と同等（最大で 7.2 μs の差）であった。ただし、90–150°の推定誤差は各方向の残差の約 2–3 倍となった。被験者 B の推定誤差は 90, 120, 240–330°では各方向の残差よりも小さい。ただし、その他の方向の推定誤差は残差よりも 10 μs 以上大きく、30, 60, 150, 180, 210°の推定誤差は残差の約 2–3.5 倍となった。被験者 C の各方向の推定誤差は 0, 90, 300°を除き、各方向の残差と同等（最大で 4.2 μs の差）であった。ただし、0, 90, 300°の推定誤差は各方向の残差よりも 10 μs 以上大きく、特に 0, 90°の推定誤差は残差の約 2 倍となった。被験者 D の各方向の推定誤差は 120, 150°を除き、各方向の残差と同等（最大で 4.4 μs の差）であった。ただし、120, 150°の推定誤差は各方向の残差の約 2–3 倍となった。

次に、各方向の全被験者の平均推定誤差（右端の Ave.）を重回帰式の残差の絶対値の平均（表 2.5 の E）と比較する。0–180°（右半面および前後方向）の平均推定誤差は、いずれの方向においても各方向の残差よりも 4.9–12.6 μs 大きい。右半面の推定誤差が残差よりも大きくなった理由については今後検討が必要である。210–330°（左半面）では、210°の推定誤差と残差の差は 1.6 μs であり、240–330°の推定誤差では各方向の残差よりも小さいことから、左半面では重回帰式と同等の精度であると言える。

全被験者の全方向の平均推定誤差は 19.2 μs であり、重回帰式の全 12 方向の残差の絶対値の平均値(16.5 μs)と 2.7 μs の差であった。この差は、アップサンプリングを施した hrir の 1 サンプル 2.604 μs と同等である。以上より、提案法はナイーブな被験者に対しても同等の精度で ITD を推定できると考えられる。ただし、一部の方向において残差よりも推定誤差が大きくなる場合があった。

次に、提案法と従来法の推定精度を比較した。上記と同じナイーブな被験者の ITD を Algazi *et al.* [61]の方法で推定した。Algazi *et al.* [61]の方法は、p1（頭の幅）、p2（高さ）、p3（奥行）の 3 箇所を用いた重回帰式から被験者の頭部半径 D[m]を導出し、式(2.4)[70]

表 2.6 ナイーブな被験者の ITD [μs]

Azimuth [deg.]	Subjects			
	A	B	C	D
0	-20.8	-49.5	-41.7	-59.9
30	281.3	268.2	296.9	294.3
60	588.5	572.9	656.3	622.4
90	697.9	721.4	752.6	742.2
120	565.1	593.8	632.8	606.8
150	270.8	294.3	312.5	278.6
180	5.2	57.3	5.2	41.7
210	-278.6	-226.6	-283.9	-281.3
240	-606.8	-562.5	-617.2	-585.9
270	-710.9	-726.6	-750.0	-731.8
300	-599.0	-619.8	-658.9	-643.2
330	-299.5	-325.5	-349.0	-320.3

表 2.7 ナイーブな被験者の頭部寸法 [mm]

	Subjects			
	A	B	C	D
p1	135	142	149	145
p2	234	237	246	252
p3	180	179	188	187
p4 _l	143	154	155	155
p4 _r	151	156	159	159
p5 _l	146	135	149	142
p5 _r	126	135	143	157
p6 _l	176	190	202	190
p6 _r	191	187	197	188
p7	369	381	398	424

表 2.8 提案法の重回帰モデルによる

推定誤差の絶対値 [μs]

Azimuth	Subjects				Ave.
	A	B	C	D	
0	21.4	24.4	35.2	15.0	24.0
30	21.4	40.9	15.1	17.0	23.6
60	20.1	53.1	11.1	12.8	24.3
90	20.0	4.1	23.2	12.2	14.9
120	49.4	2.2	15.6	38.9	26.5
150	33.5	25.4	7.6	32.5	24.8
180	0.2	44.2	20.2	18.9	20.9
210	4.2	50.0	19.8	17.0	22.7
240	26.5	9.4	0.1	20.0	14.0
270	2.5	7.6	9.0	13.3	8.1
300	5.4	17.7	29.1	7.8	15.0
330	11.8	8.0	4.9	22.0	11.7
Ave.	18.0	23.9	15.9	18.9	19.2

表 2.9 Algazi *et al.* の重回帰モデルによる

推定誤差の絶対値 [μs]

Azimuth	Subjects				Ave.
	A	B	C	D	
0	20.8	49.5	59.9	41.7	43.0
30	28.0	9.9	31.1	30.5	24.9
60	115.3	90.0	130.4	158.4	123.5
90	62.0	72.5	81.1	83.7	74.8
120	91.8	110.9	114.8	135.0	113.1
150	17.6	35.9	15.4	46.1	28.8
180	5.2	57.3	41.7	5.2	27.3
210	25.4	31.8	18.0	17.5	23.2
240	133.5	79.6	94.0	119.3	106.6
270	75.0	77.7	70.7	81.1	76.1
300	125.7	136.9	151.3	161.0	143.7
330	46.3	67.2	57.1	82.6	63.3
Ave.	62.2	68.3	72.1	80.2	70.7

表 2.10 提案法の重回帰モデルによる

方位角誤差の絶対値 [deg.]

Azimuth	Subjects				Ave.
	A	B	C	D	
0	-	-	-	-	-
30	2.4	4.9	1.6	1.8	2.7
60	2.6	7.1	1.4	1.4	3.1
90	-	0.7	4.3	2.2	2.4
120	6.7	0.3	2.0	4.3	3.3
150	4.0	2.7	0.9	3.2	2.7
180	-	-	-	-	-
210	0.5	7.1	2.2	1.9	2.9
240	3.1	1.3	0.0	2.3	1.7
270	0.0	1.4	1.7	2.4	1.4
300	0.7	2.2	3.4	0.8	1.8
330	1.3	0.7	0.5	1.9	1.1
Ave.	2.4	2.8	1.8	2.2	2.3

表 2.11 Algazi *et al.* の重回帰モデルによる

方位角誤差の絶対値 [deg.]

Azimuth	Subjects				Ave.
	A	B	C	D	
0	-	-	-	-	-
30	3.1	1.1	3.3	3.2	2.7
60	13.4	10.9	14.3	16.3	13.7
90	11.8	13.2	14.2	14.4	13.4
120	11.2	12.8	13.0	14.5	12.9
150	2.0	3.8	1.7	4.6	3.0
180	-	-	-	-	-
210	2.8	4.5	2.0	1.9	2.8
240	15.0	9.8	11.1	13.2	12.3
270	13.8	14.0	12.7	14.1	13.7
300	14.3	15.0	15.9	16.5	15.4
330	4.8	6.4	5.5	7.3	6.0
Ave.	9.2	9.2	9.4	10.6	9.6

から各方向の ITD を算出するものである.

$$ITD = \frac{D}{2c} (\sin \phi + \phi) \quad (2.4)$$

ここで, c は音速(340 m/s), ϕ は方位角[deg.]である.

Algazi *et al.* [61] の方法による ITD の推定誤差の絶対値を表 2.9 に示す. 各方向の全被験者の平均推定誤差 (右端の Ave.) は 23.2–143.7 μ s であり, 60, 120, 240, 300°では 100 μ s を超えた. いずれの方向においても, 提案法 (8.1–26.5 μ s) の方が小さい. また, 被験者ごとの全方向の平均推定誤差 (下段の Ave.)は 62.2–80.2 μ s であり, いずれの被験者においても提案法(15.9–23.9 μ s)の方が小さい. 全被験者の全方向の平均推定誤差は 70.7 μ s であり, 提案法(19.2 μ s)の方が小さい.

渡邊ら[62]の方法については, ITD を推定するために必要な回帰係数が示されていないため, 同じタイプの被験者を用いて推定精度の比較を行うことができない. ただし, 提案法における被験者ごとの平均推定誤差(15.9–23.9 μ s)は, 渡邊ら[62]の方法における方位角に対する ITD の振舞いを数式によりモデル化する際に生じる誤差 (25 μ s: 59 名の平均) よりも小さい.

次に, 提案法および Algazi *et al.* [61]の方法で推定した ITD を式(2.4)により方位角 ϕ に変換した. この方位角と元の音源の方位角の差の絶対値を求め, 方位角誤差とした. ただし, 推定した ITD が被験者本人の ITD の最大値よりも大きくなった場合は, 方位角が算出できないため計算から除外した.

方位角 30, 60, 90, 120, 150°における方向知覚の弁別閾は, 提示音の周波数に依存するものの, 最小で約 2, 3, 10, 3, 2°であることがそれぞれ示されている[63,71].

表 2.10 に提案法による方位角誤差を示す. 各方向の全被験者の平均方位角誤差 (右端の Ave.) は, 方位角 30–150° (右半面) では, それぞれ 2.7, 3.1, 2.4, 3.3, 2.7°であり方向知覚の弁別閾と同等である. 方位角 210–330° (左半面) では, それぞれ 2.9, 1.7, 1.4, 1.8, 1.1°でありいずれの方向においても方向知覚の弁別閾と同等であり, 240–330°では 2°以下となった. ただし, 各被験者の各方向の方位角誤差をみると, 一部において方向知覚の弁別閾を超える場合があり最大推定誤差は 7.1°(被験者 B の 60, 210°)であった. 各被験者の全方向の平均方位角誤差 (下段の Ave.) は 1.8–2.8°であり, 全被験者の平均方位角誤差は 2.3°であった.

Algazi *et al.* [61]の方法(表 2.11)による各方向の全被験者の平均推定誤差(右端の Ave.)

は、方位角 30–150°（右半面）ではそれぞれ 2.7, 13.7, 13.4, 12.9, 3.0°であり、60–120°では方向知覚の弁別閾を超え、提案法よりもそれぞれ 10.6, 11.0, 9.6°大きい。方位角 210–330°（左半面）ではそれぞれ 2.8, 12.3, 13.7, 15.4, 6.0°であり、210–330°では方向知覚の弁別閾を超え、提案法よりもそれぞれ 10.6, 12.3, 13.6, 4.9°大きい。各被験者の全方向の平均方位角誤差（下段の Ave.）は 9.2–10.6°であり、いずれの被験者においても提案法よりも 6.4–8.4°大きい。また、全被験者の全方向の平均方位角誤差は 9.6°であり、提案法よりも 7.3°大きい。

以上より、提案法の ITD の推定誤差は従来法よりも小さく、各方向の平均推定誤差は方向知覚の弁別閾と同等であった。したがって、提案法により高い精度で ITD を推定できることを示した。

2.3 両耳間レベル差の個人化

2.3.1 本節の目的

本節では、左右方向知覚のもう 1 つの手掛かりである ILD の個人化方法について取り上げる。頭部の前後・左右非対称性に対応するモデルを導入することで水平面内の ILD の推定精度の向上を図った。また、方位角ごとに、水平面内 12 方向の 1/3 octave band ごとの ILD を目的変数、頭部形状を説明変数とした重回帰式を構築することで方向ごとの ILD と頭部形状の相関関係の違いに対応させた。

2.3.2 従来の研究

ILD の個人化方法には、Watanabe *et al.*[64]が提案したものがあある。Watanabe *et al.* [64]は 1/3 octave band ごとの水平面内の ILD を正弦関数の和でモデル化し、その係数を頭部形状から推定する方法を提案し、4 名の被験者を用いて ILD の推定精度を検証した。水平面内 72 方向の平均推定誤差を被験者ごとに求めた結果、4 名の被験者の最大推定誤差は 1,250 Hz 以下の帯域では約 2–3 dB、それ以上の帯域では 3–8 dB であり、全ての帯域で方位角に関する ILD の弁別閾（約 1 dB[65]）を超えていた。

Watanabe *et al.*[64]の方法（以降、従来法と呼ぶ）には以下のような解決すべき課題がある。

- 1) 頭部形状の左右非対称性が考慮されていない。したがって、ILD の左右非対称性を反映できない。

- 2) 各頭部形状の回帰係数を全ての方位角で同一として扱っている。したがって、ITD と各頭部形状の相関関係の方位角による違いに対応できない。

2.3.3 両耳間レベル差の算出

a) 頭部インパルス応答の測定

ITD の推定と同様の 33 名の被験者の hrir を用いた。

b) 両耳間レベル差の算出方法

33 名の hrir から水平面内 12 方向の ILD を 1/3 octave band ごとに算出した。ILD の算出方法は以下の通りである。

- 1) 512 サンプルの hrir の後ろにゼロパディングし、48,000 サンプルでフーリエ変換する (周波数分解能 1 Hz)。
- 2) 中心周波数が 200–16,000 Hz の 1/3 octave band (20 帯域) のエネルギー E を求める。
- 3) 式(2.5)より ILD を求める。

$$ILD = 10 \log_{10} \left(\frac{E_r}{E_l} \right) \quad [\text{dB}] \quad (2.5)$$

c) 両耳間レベル差の算出結果

33 名の ILD を図 2.4 に示す。音源の方位角は正面方向を 0° として時計回りに表した。正の値は右耳の音圧レベルが大きいことを表す。

ILD の絶対値の最大値は、周波数が高くなるにつれて大きくなる傾向があり、500 Hz 以下では約 5 dB、630–1,000 Hz では約 10 dB であった。1,250 Hz 以上の帯域では 20 dB を超え、6,300 Hz 以上では 40 dB を超えた。いずれの帯域においても $0, 180^\circ$ (正面と後方) で 0 dB とならない場合があるが、これは頭部の左右非対称性が影響していると考えられる。

また、方位角に対する ILD の振舞いに不連続性が観察される。例えば、2000 Hz では、多くの被験者において $90, 270^\circ$ (真横) の ILD はその前後の ILD よりも小さくなっている。提案法では、方位角に対する ILD の振舞いをモデル化せずに、方向ごとの ILD を推定するため、ILD の振舞いの不連続性にも対応できると考えられる。

ILD の個人差も周波数が高くなるにつれて大きくなる傾向がある。1,250 Hz 以下ではいずれの方向においても 5–10 dB の個人差がある。1,600–10,000 Hz では、いずれの方

向においても 10–20 dB の個人差がある。12,500, 16,000 Hz ではいずれの方向においても 20 dB 以上の個人差があり、60° では 40 dB に及んだ。いずれの帯域においても、90° の ILD の分布範囲は 60, 120° のそれらと互いに重なり、270° の分布範囲も 240, 300° のそれらと互いに重なっていた。特に、4,000 Hz では 30–150° の分布範囲は互いに重なっており、210–330° においても同様である。これらの帯域の ILD には 30° 分以上の個人差が存在する。

図 2.5 に一例として、2 名の被験者の 4000 Hz の ILD を示す。被験者 2 の 90° の ILD は、被験者 1 の 60, 120° とほぼ一致している。方位角知覚の弁別閾が前方では 1°、側方では約 10° であることから[63]、他の帯域においても、他人の ILD を用いると方位角の違いが知覚されると考えられる。

次に、ILD の前後・左右非対称性についてみる。全被験者および全方向について ILD の左右・前後非対称性の平均値および最大値を帯域ごとに求めた。その結果を表 2.12 に示す。

前後非対称性の平均値(Ave.)についてみると、周波数が高くなるにつれて大きくなる傾向があり、全帯域の平均は 3.0 dB であり、800 Hz 以上の帯域では ILD の弁別閾(約 1 dB)よりも大きい。左右非対称性の平均値についても同様の傾向があり、全帯域の平均は 2.1 dB であり、1250 Hz 以上の帯域では ILD の弁別閾(約 1 dB)よりも大きい。また、各帯域における全被験者の最大値(Max)は、前後・左右非対称性ともに、いずれの帯域においても ILD の弁別閾(約 1 dB)よりも大きい。以上の結果は、ILD の前後・左右非対称性を考慮することで ILD の推定精度が向上する可能性を示唆している。

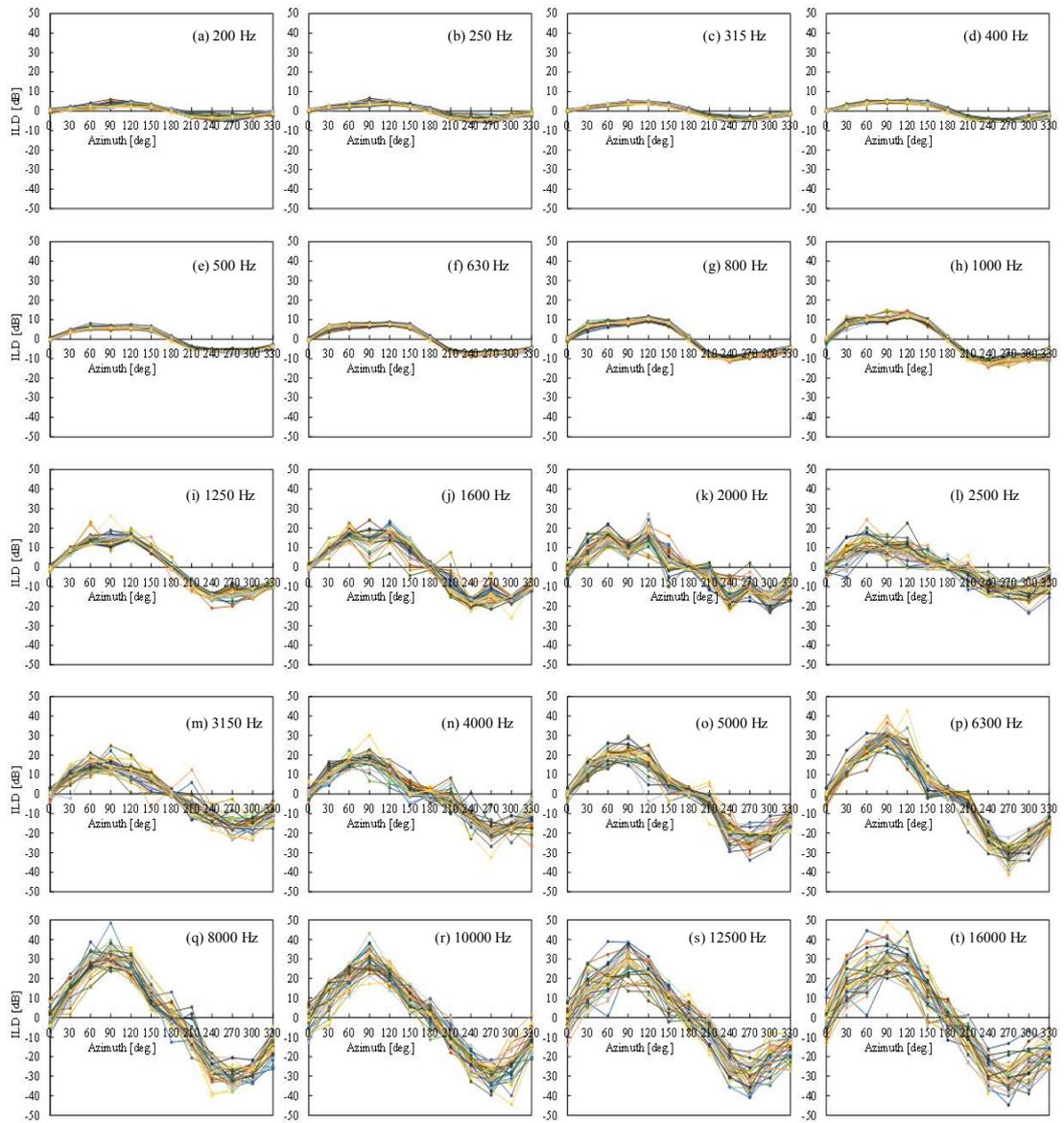


図 2.4 33名の水平面でのILD

表 2.12 ILD の左右・前後非対称性の平均値と最大値

Center frequencies of 1/3 octave band [Hz]	Asymmetry of the ILD [dB]			
	Average		Max	
	Front-Back	Left-Right	Front-Back	Left-Right
200	0.9	0.8	3.4	4.4
250	1.0	0.7	3.3	4.0
315	0.8	0.5	2.4	3.1
400	0.7	0.5	2.9	3.3
500	0.6	0.5	2.7	3.3
630	0.9	0.6	2.8	2.8
800	1.6	0.8	5.3	3.7
1000	1.5	0.9	5.8	5.4
1250	1.5	1.3	7.9	12.6
1600	2.3	2.1	13.0	11.3
2000	4.1	2.7	19.4	20.1
2500	3.6	2.4	14.8	12.7
3150	3.5	2.5	23.4	12.2
4000	5.2	2.9	25.8	15.3
5000	4.8	3.0	22.8	17.1
6300	5.1	3.4	23.5	15.0
8000	4.5	3.6	22.0	16.7
10000	4.5	4.3	20.8	20.6
12500	5.5	4.7	21.6	24.1
16000	6.4	4.9	26.8	22.7
Ave.	3.0	2.1	13.5	11.5

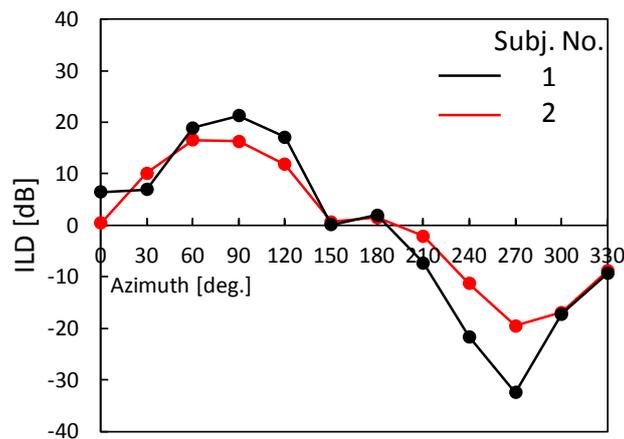


図 2.5 2名の水平面での ILD の比較

2.3.4 頭部形状の計測

頭部表面を通る音波の伝播経路の個人差が ILD の個人差に対応すると考え、ITD の個人化と同様の頭部に関する部位だけを用いた頭部モデルを導入し、ILD を求めた 33 名の頭部形状を計測した。

Watanabe *et al.* [64]は ILD を推定するための頭部モデルとして、頭部の幅、長さ、奥行、鼻の高さ、長さ、奥行、耳介の幅、長さ、肩幅、頭耳長、頭耳高の 11 箇所を用いている。ただし、頭部の前後・左右非対称性は考慮されていない。

頭部モデルは ITD の推定と同様のものを用いた (図 2.6)。Watanabe *et al.* [64]の頭部モデルのうち、鼻および耳介に関するものを除き、全ての部位を含めている。このうち、 $p4_{l,r}$, $p5_{l,r}$ および $p6_{l,r}$ は、Watanabe *et al.* [64]の頭耳長および頭耳高に相当する部位であるが、頭部の前後・左右非対称性に対応できるように前後・左右に分割した。

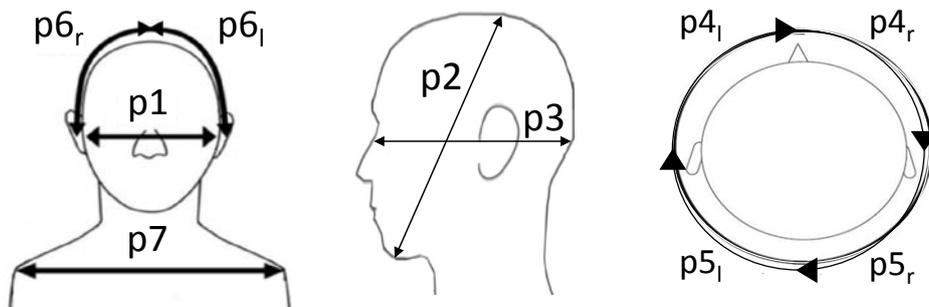


図 2.6 ILD の推定に用いた頭部モデル

2.3.5 単回帰分析

各方向における ILD と各頭部形状の単相関係数を求めた。一例として表 2.13–2.14 に 250, 1000, 4000 Hz の結果を示す。表中の網掛けは有意水準 10%で相関関係が認められることを表す($0.29 \leq |r|$)。

250 Hz (表 2.13) では、 $p7$ は 3 方向、 $p3, p4_{l,r}$ は 2 方向、 $p1, p2, p5_r$ は 1 方向で有意な相関が認められた。ただし、12 方向中 6 方向で有意な相関が認められる部位が存在しなかった。

1000 Hz (表 2.14) では、 $p1$ は後方、 $p2$ は側方、 $p3$ は右側方で有意な相関が認められる傾向にある。 $p4_{l,r}$ は 120, 240° (側方から 30°後方) で比較的高い相関が得られた。 $p5_{l,r}$ は 0°もしくは 150°で有意な相関が認められた。 $p6_l$ は有意な相関は認められなかった。 $p6_r$ は 120, 240° (側方から 30°後方) で相関が認められた。 $p7$ は 30, 60, 120, 240, 300, 330°

の 6 方向において有意な相関が認められた。以上より、250 Hz に比べて多くの箇所では有意な相関がみられた。有意な相関が認められない方向は存在しなかった。

4000 Hz (表 2.15) では、 p_2 , p_3 , $p_{4_{lr}}$, p_{6_r} , p_7 は 60° 方向でのみ有意な相関が認められた。 p_1 は $60, 270^\circ$, p_{5_r} は $90, 300^\circ$, p_{6_l} は 0° で相関が認められた。ただし、12 方向中 7 方向で有意な相関が認められる部位が存在しなかった。

以上より、有意な相関となる部位が存在しない方向があることから、単回帰分析では ILD を精度良く推定することは困難であると考えられる。

表 2.13 頭部形状と各方向の ILD の単相関係数 r (250 Hz).

：10%有意($0.29 \leq |r|$)

Azm.	p1	p2	p3	p4 _l	p4 _r	p5 _l	p5 _r	p6 _l	p6 _r	p7
0	-0.14	-0.30	-0.02	-0.18	-0.29	0.03	0.28	-0.23	0.18	-0.40
30	-0.01	-0.20	0.14	-0.22	-0.15	0.09	0.33	-0.20	0.07	-0.29
60	0.25	-0.10	0.26	0.07	0.26	0.07	0.18	0.05	0.17	-0.24
90	0.06	0.11	0.16	-0.04	0.01	0.00	0.20	0.02	0.11	0.06
120	0.30	-0.02	0.32	0.08	0.27	-0.01	0.27	-0.03	0.25	0.03
150	0.19	-0.04	0.15	0.09	0.17	-0.22	0.00	0.07	0.19	-0.16
180	0.09	-0.01	-0.06	-0.23	0.03	-0.25	-0.08	0.20	-0.05	-0.19
210	-0.16	-0.03	-0.08	-0.28	-0.13	-0.15	0.01	0.08	-0.15	-0.22
240	-0.07	-0.17	-0.12	-0.34	-0.12	-0.01	0.04	0.09	-0.22	-0.22
270	0.04	-0.19	-0.21	-0.32	-0.04	-0.04	-0.04	0.16	0.09	-0.29
300	-0.18	-0.03	-0.33	-0.27	-0.38	0.15	-0.18	0.13	-0.12	-0.28
330	0.07	-0.08	0.06	-0.05	-0.03	0.25	0.08	-0.08	0.15	-0.04

表 2.14 頭部形状と各方向の ILD の単相関係数 r (1000 Hz)

：10%有意($0.29 \leq |r|$)

Azm.	p1	p2	p3	p4 _l	p4 _r	p5 _l	p5 _r	p6 _l	p6 _r	p7
0	0.09	-0.18	0.18	-0.01	0.01	0.42	0.19	-0.05	0.13	-0.15
30	0.21	0.38	0.32	0.21	0.25	0.10	0.11	-0.02	0.23	0.57
60	0.20	0.47	0.44	0.29	0.37	0.18	0.23	-0.08	0.24	0.58
90	0.18	0.36	0.24	0.25	0.25	0.04	-0.05	0.05	0.02	0.14
120	0.31	0.61	0.42	0.47	0.51	-0.04	0.01	0.13	0.38	0.38
150	0.04	0.08	0.39	0.05	-0.05	0.49	0.32	-0.08	0.26	0.02
180	0.36	0.10	0.23	0.15	0.32	-0.17	0.11	0.10	0.19	0.04
210	0.35	0.19	0.05	0.23	0.33	-0.26	-0.03	0.24	0.14	0.08
240	-0.15	-0.50	-0.24	-0.55	-0.33	0.15	-0.10	-0.14	-0.34	-0.49
270	-0.07	-0.29	-0.01	-0.24	-0.01	0.08	0.04	-0.17	0.04	-0.08
300	0.05	-0.36	-0.20	-0.24	0.01	-0.13	-0.08	-0.07	-0.01	-0.31
330	-0.19	-0.36	-0.27	-0.34	-0.31	0.26	-0.09	0.02	-0.08	-0.62

表 2.15 頭部形状と各方向の ILD の単相関係数 r (4000 Hz).

：10%有意($0.29 \leq |r|$)

Azm.	p1	p2	p3	p4 _l	p4 _r	p5 _l	p5 _r	p6 _l	p6 _r	p7
0	-0.19	-0.24	0.01	-0.11	-0.09	0.14	-0.07	-0.41	0.00	-0.16
30	0.20	0.23	0.02	0.10	-0.16	-0.06	0.16	0.19	-0.19	0.13
60	0.29	0.46	0.34	0.38	0.32	0.06	0.12	0.03	0.35	0.37
90	0.07	-0.09	-0.08	0.00	0.12	-0.08	-0.32	0.23	0.05	-0.28
120	0.14	-0.04	0.03	0.15	-0.04	-0.01	-0.08	0.17	0.17	-0.16
150	0.19	0.24	0.17	0.18	0.09	-0.23	-0.08	0.13	-0.12	0.09
180	0.01	-0.26	-0.05	-0.04	0.10	0.15	-0.18	-0.17	0.03	-0.08
210	0.05	-0.19	0.16	-0.12	0.12	0.15	0.16	-0.04	-0.09	-0.13
240	0.05	-0.28	-0.11	-0.17	-0.09	0.01	0.22	0.05	0.03	-0.03
270	0.29	0.23	0.22	0.17	0.19	-0.02	0.27	0.01	0.08	0.26
300	0.09	-0.24	-0.04	0.12	-0.07	0.11	0.29	-0.08	-0.06	0.03
330	-0.07	-0.11	0.03	-0.06	0.16	0.08	-0.10	-0.24	0.19	0.04

2.3.6 重回帰分析

10箇所の頭部形状を説明変数, 各方向のILDを目的変数とした重回帰分析を行った.

$$ILD(s, \varphi, f) = a_1(\varphi, f)p_1 + a_2(\varphi, f)p_2 + \dots + a_7(\varphi, f)p_7 + \beta(\varphi, f) \quad [\text{dB}] \quad (2.6)$$

ここで, s は被験者, φ は方位角, f は周波数である.

なお, 全ての説明変数の VIF (Variance Inflation Factor) が 10 未満となったことから, 説明変数間の多重共線性はないとみなせる[69].

表 2.16 に重相関係数 r を示す. 重相関係数 r は, 0.31–0.80 であり, いずれも有意水準 10% で相関関係が認められた ($0.29 \leq |r|$). また, 250 Hz の 90°, 1250 Hz の 30°, 10000 Hz の 210° を除き, 有意水準 5% で相関関係が認められた ($0.34 \leq |r|$). 全体の平均相関係数は 0.58 であった. 各方向における平均値 (右端の Ave.) は 0.53–0.64 であり, 30, 60° で僅かに高いが方向間の差は小さい. 各周波数帯域における平均値 (下段の Ave.) は, 0.49–0.66 であり, 1250 Hz でやや低い.

表 2.17 に重回帰係数の全帯域の平均値を示す. 各方向の回帰係数をみると, 0, 60, 120, 150° では α_3 (頭の奥行), 240, 300, 330° では α_2 (頭の高さ), 180, 210° では α_1 (頭の幅), 30, 90° では α_4 (頭周の左前), 270° では α_4 (頭周の右前) が最も大きい値をとった.

表 2.18 に危険率 p を示す. 網掛け部分は p 値が 0.10 未満となったことを表す. 全体の平均危険率 p は 0.42 であった. 方向ごとにみると, 180° では全ての帯域で 0.10 以上となったが, その他の方向では 1–7 帯域で 0.10 未満となった. 各方向の平均値 (右端の Ave.) は, 0.27–0.55 であり, 右側方で低くなる傾向がある. 周波数ごとにみると, 危険率 p は 1250, 1600, 3150 Hz では全ての方向で 0.10 以上となったが, その他の周波数では 1–5 方向において 0.10 未満となった. 各周波数の平均値 (下段の Ave.) は, 0.22–0.68 であり, 1250 Hz でやや高い.

表 2.19 に残差の絶対値 E を示す. 全体の平均残差は 1.67 dB であった. 各方向における平均値 (右端の Ave.) は 1.16–1.87 dB であり, 0, 180° で小さい傾向にある. その他の方向においては方向間の差は小さい. 各周波数帯域における平均値 (下段の Ave.) は, 0.29–4.13 dB であり, 周波数が高くなるにつれて残差が大きくなる傾向がある. これは Watanabe *et al.*[64] と同様の傾向である. 1000 Hz 以下では方向知覚の弁別閾 (約 1 dB[62]) 以内となった.

以上より, 1000 Hz 以下では方向知覚の弁別閾内の精度で ILD を推定できる可能性が示唆される.

表 2.16 重回帰分析の結果： 重相関係数 r

Azm.	center frequencies of 1/3 octave bands [Hz]																				Ave.
[deg.]	200	250	315	400	500	630	800	1000	1250	1600	2000	2500	3150	4000	5000	6300	8000	10000	12500	16000	
0	0.64	0.78	0.64	0.52	0.45	0.50	0.62	0.59	0.52	0.54	0.72	0.74	0.52	0.57	0.53	0.78	0.54	0.65	0.53	0.39	0.59
30	0.70	0.68	0.67	0.62	0.68	0.55	0.80	0.59	0.31	0.53	0.76	0.37	0.65	0.70	0.69	0.42	0.63	0.50	0.74	0.67	0.61
60	0.79	0.67	0.71	0.80	0.75	0.59	0.75	0.66	0.47	0.56	0.52	0.59	0.54	0.54	0.61	0.77	0.53	0.61	0.65	0.74	0.64
90	0.38	0.32	0.63	0.75	0.66	0.59	0.62	0.48	0.55	0.44	0.51	0.48	0.66	0.54	0.67	0.57	0.62	0.69	0.62	0.74	0.58
120	0.53	0.57	0.71	0.73	0.68	0.74	0.75	0.72	0.41	0.50	0.64	0.40	0.48	0.55	0.45	0.41	0.61	0.54	0.74	0.60	0.59
150	0.57	0.57	0.65	0.68	0.64	0.65	0.63	0.68	0.49	0.61	0.67	0.44	0.55	0.59	0.59	0.70	0.72	0.45	0.51	0.53	0.60
180	0.58	0.64	0.66	0.65	0.61	0.57	0.55	0.52	0.53	0.60	0.61	0.53	0.55	0.49	0.51	0.51	0.52	0.49	0.44	0.68	0.56
210	0.50	0.54	0.58	0.67	0.56	0.35	0.40	0.52	0.42	0.63	0.70	0.53	0.56	0.54	0.62	0.46	0.62	0.32	0.50	0.54	0.53
240	0.51	0.53	0.68	0.70	0.63	0.53	0.58	0.72	0.51	0.58	0.57	0.66	0.50	0.52	0.51	0.49	0.72	0.59	0.54	0.45	0.58
270	0.56	0.63	0.73	0.73	0.67	0.49	0.52	0.46	0.53	0.61	0.56	0.69	0.65	0.42	0.78	0.55	0.40	0.36	0.74	0.77	0.59
300	0.59	0.65	0.57	0.55	0.67	0.55	0.48	0.52	0.64	0.53	0.71	0.57	0.58	0.62	0.45	0.46	0.59	0.52	0.63	0.80	0.58
330	0.60	0.34	0.36	0.49	0.61	0.71	0.57	0.78	0.50	0.47	0.60	0.39	0.34	0.51	0.46	0.43	0.71	0.49	0.71	0.64	0.54
Ave.	0.58	0.58	0.63	0.66	0.63	0.57	0.61	0.60	0.49	0.55	0.63	0.53	0.55	0.55	0.57	0.55	0.60	0.52	0.61	0.63	0.58

表 2.17 重回帰分析の結果： 重回帰係数の全帯域の平均値（頭部形状の単位は mm）

Azm [deg.]	Regression coefficient ($\times 10^2$)										
	α_1	α_2	α_3	α_{4_l}	α_{4_r}	α_{5_l}	α_{5_r}	α_{6_l}	α_{6_r}	α_7	b
0	2.40	-4.65	9.48	-2.36	-2.61	1.02	-3.72	-0.85	4.15	0.44	-656.06
30	3.15	3.15	7.09	-7.19	-0.52	-1.94	-0.05	0.01	5.35	0.19	-1375.97
60	6.93	5.65	9.17	-8.06	-1.98	-4.37	-2.91	-0.12	6.40	-0.96	-1014.93
90	4.40	5.59	3.43	-13.60	1.33	-4.27	-2.21	2.65	9.64	-0.61	-555.28
120	-0.14	-3.56	16.28	-7.22	-4.10	-5.09	-3.74	5.11	7.84	0.68	-581.92
150	-5.20	2.94	8.81	-6.11	0.55	-0.34	-2.10	0.69	8.13	-0.53	-1330.10
180	9.95	-2.93	5.47	-4.09	-3.27	0.63	-3.37	-1.15	-0.81	0.81	-67.86
210	6.28	-0.30	-4.07	-1.48	5.52	1.94	1.90	-0.21	-1.79	0.38	-1616.90
240	0.01	-6.50	-2.78	-3.48	3.47	-0.90	3.03	2.12	-1.36	2.44	-683.93
270	-2.34	-6.20	-5.64	-0.55	9.16	6.81	2.54	-1.23	-4.97	0.96	-503.77
300	1.43	-7.90	-2.51	6.79	-0.91	5.54	-1.54	-2.88	-0.37	0.76	-287.99
330	5.20	-6.59	4.44	0.29	-1.53	1.40	-3.62	-2.65	2.50	-0.12	-221.68

表 2.18 重回帰分析の結果： 危険率 p () : $p < 0.10$

Azm.	center frequencies of 1/3 octave bands [Hz]																				Ave.
[deg.]	200	250	315	400	500	630	800	1000	1250	1600	2000	2500	3150	4000	5000	6300	8000	10000	12500	16000	
0	0.20	0.01	0.20	0.61	0.82	0.69	0.27	0.36	0.62	0.55	0.04	0.03	0.61	0.43	0.59	0.01	0.53	0.16	0.57	0.94	0.41
30	0.07	0.10	0.13	0.27	0.09	0.49	0.004	0.34	0.99	0.59	0.02	0.96	0.16	0.07	0.09	0.88	0.23	0.67	0.02	0.12	0.32
60	0.01	0.11	0.06	0.004	0.02	0.38	0.02	0.15	0.78	0.48	0.62	0.35	0.55	0.56	0.28	0.01	0.58	0.29	0.16	0.03	0.27
90	0.95	0.99	0.23	0.02	0.14	0.37	0.27	0.74	0.52	0.86	0.64	0.76	0.14	0.55	0.13	0.42	0.25	0.08	0.25	0.03	0.42
120	0.57	0.45	0.05	0.03	0.11	0.03	0.02	0.05	0.91	0.68	0.20	0.92	0.75	0.49	0.83	0.90	0.30	0.53	0.02	0.31	0.41
150	0.42	0.42	0.16	0.11	0.18	0.18	0.22	0.10	0.73	0.29	0.13	0.85	0.49	0.37	0.36	0.07	0.05	0.83	0.64	0.57	0.36
180	0.38	0.19	0.14	0.16	0.29	0.44	0.51	0.61	0.59	0.31	0.28	0.59	0.50	0.73	0.66	0.65	0.60	0.71	0.86	0.11	0.47
210	0.68	0.55	0.39	0.13	0.48	0.97	0.92	0.62	0.90	0.22	0.06	0.58	0.45	0.56	0.25	0.79	0.25	0.99	0.69	0.55	0.55
240	0.67	0.58	0.09	0.06	0.23	0.59	0.40	0.04	0.64	0.41	0.42	0.15	0.69	0.63	0.65	0.70	0.04	0.34	0.55	0.83	0.44
270	0.46	0.23	0.03	0.03	0.13	0.71	0.61	0.80	0.59	0.29	0.47	0.09	0.17	0.90	0.01	0.51	0.93	0.97	0.03	0.01	0.40
300	0.35	0.16	0.45	0.50	0.12	0.52	0.75	0.60	0.20	0.57	0.06	0.42	0.40	0.26	0.83	0.81	0.37	0.61	0.21	0.004	0.41
330	0.33	0.98	0.96	0.73	0.27	0.06	0.42	0.01	0.68	0.77	0.32	0.93	0.98	0.67	0.81	0.87	0.05	0.73	0.05	0.19	0.54
Ave.	0.42	0.40	0.24	0.22	0.24	0.45	0.37	0.37	0.68	0.50	0.27	0.55	0.49	0.52	0.46	0.55	0.35	0.58	0.34	0.31	0.42

表 2.19 重回帰分析の結果： 残差の絶対値の平均 E (dB)

Azm. [deg.]	center frequencies of 1/3 octave bands [Hz]																				Ave.
	200	250	315	400	500	630	800	1000	1250	1600	2000	2500	3150	4000	5000	6300	8000	10000	12500	16000	
0	0.27	0.23	0.21	0.24	0.28	0.33	0.47	0.58	0.62	0.57	1.32	1.47	1.09	1.48	1.06	1.20	3.03	2.94	3.62	4.39	1.27
30	0.36	0.34	0.29	0.32	0.38	0.53	0.65	1.00	0.73	1.28	2.09	3.20	1.96	1.99	1.62	1.61	2.55	3.79	3.72	4.35	1.64
60	0.40	0.43	0.26	0.29	0.33	0.45	0.52	0.66	1.53	1.39	2.26	2.35	1.97	1.57	2.11	1.47	3.07	3.16	4.06	4.93	1.66
90	0.76	0.67	0.33	0.27	0.30	0.43	0.49	0.89	1.92	3.10	1.45	2.02	1.88	2.51	2.09	2.73	3.27	3.04	4.69	3.71	1.83
120	0.63	0.53	0.24	0.25	0.35	0.38	0.46	0.50	0.81	2.10	2.49	2.89	2.38	2.27	3.40	4.03	2.97	2.72	2.92	5.16	1.87
150	0.45	0.37	0.26	0.30	0.34	0.35	0.52	0.52	1.01	2.78	3.25	2.00	2.66	2.29	1.95	2.02	2.28	3.11	3.54	5.44	1.77
180	0.33	0.34	0.31	0.36	0.44	0.48	0.52	0.50	1.01	0.88	0.70	0.94	0.74	1.27	0.94	0.97	2.73	3.32	3.55	2.94	1.16
210	0.49	0.48	0.38	0.35	0.36	0.47	0.68	0.83	0.96	2.68	2.08	2.33	3.06	2.60	2.45	2.85	2.92	3.25	3.63	3.85	1.84
240	0.72	0.68	0.33	0.26	0.29	0.44	0.51	0.53	0.94	1.47	3.58	1.82	1.89	2.66	3.31	2.87	2.92	2.88	4.15	4.31	1.83
270	0.60	0.58	0.27	0.28	0.38	0.46	0.47	0.92	1.79	2.61	2.08	1.33	2.28	2.68	2.12	3.15	3.02	3.60	3.50	3.26	1.77
300	0.47	0.39	0.31	0.38	0.37	0.41	0.53	0.67	0.93	1.18	1.80	2.44	2.24	1.42	2.41	2.81	2.38	5.19	3.93	3.69	1.70
330	0.43	0.45	0.32	0.39	0.41	0.33	0.54	0.74	0.90	1.23	2.28	2.91	2.52	3.63	2.33	1.77	3.07	3.72	2.95	3.60	1.72
Ave.	0.49	0.46	0.29	0.31	0.35	0.42	0.53	0.70	1.10	1.77	2.12	2.14	2.05	2.20	2.15	2.29	2.85	3.39	3.69	4.13	1.67

2.3.7 ナイーブな被験者による両耳間レベル差の推定精度の検証

2.3.6 項で求めた重回帰モデルの推定精度を重回帰分析に含めていない被験者を用いて検証した。被験者は20代の女性2名(A, B)と男性2名(C, D)の4名である。この4名のILDと頭部形状を求めた。なお、この4名はITDの推定精度の検証に用いた被験者と同じである。

重回帰式を用いて頭部形状から12方向、20帯域のILDを推定した。図2.7に推定誤差の絶対値を示す。推定誤差は周波数が高くなると大きくなる傾向がある。200–800 Hzでは、いずれの被験者においてもほぼ全ての方向で1 dB以下となった。ただし、200, 250 Hzの90°において、被験者B, Cで3 dB程度の誤差が生じた。1000 Hzでは、いずれの被験者においても側方で4 dB程度の誤差が生じた。1250–4000 Hzでは、いずれの被験者においても5–10 dBの誤差が生じる場合がある。5000 Hz以上ではいずれの被験者においても10 dBを超える誤差となる場合がある。

図2.8に被験者ごとに12方向の推定誤差の絶対値を平均した結果を示す。いずれの被験者においても、周波数が高くなるにつれて推定誤差が大きくなる傾向がある。200–800 Hzでは、いずれの被験者においてもILDの弁別閾(約1 dB[65])以内となった。それ以上の帯域では、4名の被験者の最大平均誤差は2–9 dBとなり、ILDの弁別閾を超える誤差が生じた。特に、1250–2000 Hz付近で推定誤差が大きくなる傾向がみられ、10000 Hz以上の帯域では被験者A, Bにおいて推定誤差が大きい(6–9 dB)。

以上より、800 Hz以下では4名のナイーブな被験者に対してもILDを弁別閾内(約1 dB[65])の精度で推定可能であった。

次に従来法との比較を行う。Watanabe *et al.* [64]は本研究と同様に4名のナイーブな被験者を用いて、1/3 octave bandごとのILDの推定精度を検証した。被験者ごとに、水平面内72方向の推定誤差を平均した結果を示している。本研究とは異なる被験者であり、誤差を平均した方向数も異なるため厳密な比較は行えないが、1000 Hz以上の帯域においては、Watanabe *et al.* [64]の4名の被験者の最大推定誤差は3–8 dBであるのに対し、提案法では2–9 dBであった。特に、提案法の1250–2000, 12500, 16000 Hzの推定誤差は、のそれと比べて1 dB程度大きい。この要因としてWatanabe *et al.* [64]がILDの推定に用いていた鼻に関するパラメータを本研究では用いていないことが考えられる。一方、800 Hz以下の帯域では、Watanabe *et al.* [64]の4名の被験者の最大推定誤差は2–3 dBとなりILDの弁別閾(約1 dB[65])を超えていたが、提案法の4名の最大推定誤差

は弁別閾内となった。

以上より、800 Hz 以下では、ILD の推定誤差は弁別閾内となった。頭部の左右非対称性を考慮することで 800 Hz 以下の推定精度が向上した可能性が示された。しかし、1000 Hz 以上の帯域では、提案法の推定精度が従来法のそれに比べて改善されなかった。この要因として、提案法では、従来法で用いられている鼻と耳介に関するパラメータを除外していたことが挙げられ、今後の検討の余地を残している。

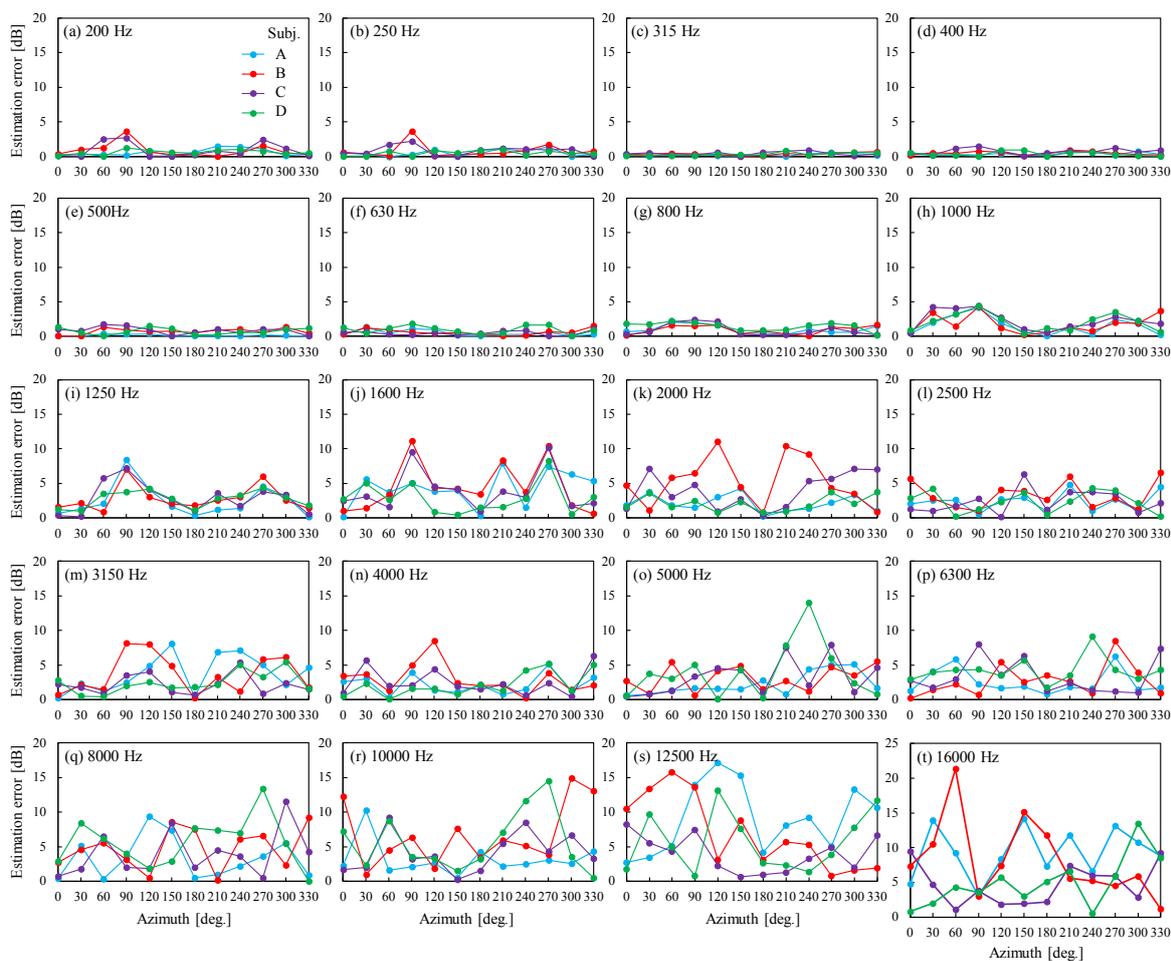


図 2.7 提案法の重回帰モデルによる推定誤差の絶対値

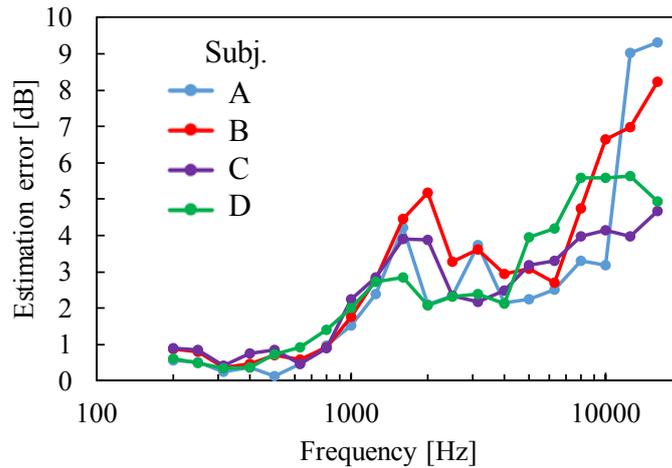


図 2.8 周波数帯域ごとの平均推定誤差

2.4 結論

本研究では，受聴者の頭部形状から ITD および ILD を推定するため，頭部の前後・左右非対称性に対応した頭部モデルを導入した．また，方位角ごとに，水平面内 12 方向の ITD および 1/3 octave band ごとの ILD を目的変数，頭部形状を説明変数とした重回帰式を構築することで方向ごとの ITD および ILD と頭部形状の相関関係の違いに対応させた．推定精度の目標値である方向知覚の弁別閾，および従来法との比較を行った．

- 1) 各方向の ITD を目的変数，10 箇所頭部形状を説明変数とした重回帰分析を行った．その結果，重相関係数は 0.34–0.79 であり，全ての方向において 5% で有意な相関関係が認められた．
- 2) 4 名のナイーブな被験者を用いて，提案法による ITD の推定精度を検証した．その結果，各方向の平均方位角誤差は 1.1–3.3° であり方向知覚の弁別閾と同等であった．ただし，各被験者の各方向の方位角誤差をみると，一部において方向知覚の弁別閾を超える場合があり，最大推定誤差は 7.1°（被験者 B の 60, 210°）であった．また，提案法の平均方位角誤差は 2.3° であり，従来法を用いた場合の平均方位角誤差よりも 7.3° 小さかった．
- 3) 10 箇所頭部形状を説明変数，1/3 octave band (250–16000 Hz) の ILD を目的変数とした重回帰分析を行った．その結果，重相関係数の平均値は 0.58 であり，残差の絶対値の平均値は 1.67 dB であった．
- 4) 4 名のナイーブな被験者を用いて重回帰式の ILD の推定精度を検証した．その結果，800 Hz 以下では ILD を弁別閾内（約 1 dB）の精度で推定可能であることを示した．

ただし、1000 Hz 以上では、推定誤差は 2–9 dB で従来法と同等であり、高い精度で推定することはできなかつたため、今後の検討の余地を残している。

- 5) 提案法による ILD の個人化に要する頭部形状の計測時間は、約 3 分であつた。また、特別な設備および計測者の訓練を必要しない。

以上より、頭部の左右非対称性を考慮し、方向ごとに重回帰モデルを構築した結果、ITD および 800 Hz 以下の ILD は弁別閾内の精度で推定できることを示した。しかし、1000 Hz 以上の帯域では、提案法の推定精度が従来法のそれに比べて改善されなかつた。この要因として、提案法では、従来法で用いられている鼻と耳介に関するパラメータを除外していたことが挙げられ、検討の余地を残した。

第3章 スペクトラルキューの個人化

3.1 本章の目的

第2章では、水平面内の両耳間差キューの個人化について検討した。本章では、前後上下方向知覚の手掛かりであるスペクトラルキューについて採り上げる。第1章で述べように、他人の HRTF を用いると、音像の上昇、前後誤判定、頭内定位などの現象がしばしば発生する。スペクトラルキューの個人化については、従来、多数の方法が提案されているが、受聴者本人の HRTF と同等の定位精度は得られていない。本研究では、HRTF に含まれるピークやノッチのうち、特に N1, N2 がスペクトラルキューとして重要な役割を果たすことに着目し、N1, N2 を各個人に適合させることで、高い音像制御が実現できるのではないかと考えた。

本章では、受聴者の頭部形状から N1, N2 の周波数を推定し、その周波数に近い N1, N2 周波数を持つ HRTF をデータベースから選出することで HRTF の個人化を行う方法を提案した。この方法により選出した HRTF を *best-matching HRTF* と呼ぶ。本章では、*best-matching HRTF* が本人 HRTF と同等の定位精度を有するか否かにより、提案法の有効性を検証した。

3.2 従来の研究

スペクトラルキューの個人化方法に関する従来法には以下のような課題がある。

- 1) 個人に適合する HRTF を生成する方法
- 2) 既存の HRTF を個人に適合させる方法
- 3) データベースから適合する HRTF を選出する方法

以降、それぞれの方法について概説する。

1) 個人に適合する HRTF を生成する方法

Kistler and Wightman[55]は、PCA(principal component analysis)によって、HRTF をいくつかの主成分に分解できることを示し、最初の5個の主成分で HRTF の振幅スペクトルの90%が再現できるとした。PCA を利用して、受聴者の耳介形状から各主成分の係数を推定し、個人化 HRTF を合成する試みがある[56]。しかし、いずれの文献においても、この方法により個人化した HRTF の音像定位精度は示されておらず、その有効性は明ら

かではない。また、寄与率の高い主成分がスペクトラルキューに対応しているとは限らないため、スペクトラルキューとして重要なノッチを再現するには不十分であり、ノッチを再現するためには主成分が 15–20 程度必要であるとする計算例もある。

2) 既存の HRTF を個人に適合させる方法

Middlebrooks[57-59]は、標準的な HRTF を受聴者の耳介形状に応じて周波数軸上で伸縮する方法を提案した。彼は、代表的な HRTF の方向依存性(DTF: directional transfer function)を周波数軸上で適切に伸縮することで、受聴者本人の DTF に近似することが可能であり、音像定位実験により、他人の DTF を適切に伸縮した DTF の quadrant error (前後もしくは上下方向における 90° 以上の誤差のことで、前後誤判定率、上下誤判定率を足し合わせたようなもの) が 14.7%であり、被験者本人の DTF の quadrant error 15.6%と同等であることを示した。ただし、受聴者に適合した伸縮率を見いだすには、1ブロック 20分を要する試聴実験を1から3ブロック実施する必要があるという問題があり、適切な伸縮率を受聴者の耳介形状から導くことには成功していない。

3) データベースから適合する HRTF を選出する方法

この方法の代表的な取り組みとして、Zotkin *et al.* [60]が提案している。このアプローチは、耳介の形状が似ていれば、HRTF も似ているという仮説に支えられている。彼らは、8人の受聴者について耳介部位 7箇所を計測し、その計測値に最も近い耳介の HRTF をデータベースから選出した。この HRTF を用いて前半球面を目標方向とした音像定位実験を行ったが、その音像定位精度はダミーヘッドの HRTF と比較してわずかに 1.9° 改善されただけであった。この方法の問題点として、すべての耳介部位の個人差を同じ重みで扱っていることが挙げられる。前述のように、HRTF のピークは耳甲介腔の深さ方向のモード、耳介の上下方向のモードに由来し、ノッチは外耳道入口付近に生じた節に由来することから、これらの現象に関連する部位に着目して耳介形状の類似度を評価する必要がある。

このような従来法の問題点として、耳介形状の個人差を全ての部位で同じ重みで扱っていること、HRTF のスペクトル構造の個人差を全ての周波数帯域で同じ重みで扱っていることが挙げられる。スペクトラルキューとして N1,N2 が重要な役割を果たしていること、および N1,N2 は外耳道入口における節に由来することから、これらの情報に重

みを付ける方法が有効であると考えられる。

3.3 頭部伝達関数の測定

HRTF の測定は無響室で行った。測定方向は水平面内 12 方向 (30°間隔), スピーカから頭部中心位置までの距離は 1.2 m である。実験システムは, ノート PC, オーディオインターフェース(RME FireFaceUFX), スピーカ(FOSTEX FE83E), アンプ(Marantz PM4001), A/D コンバータ(Roland EDIROL M-10MX), 耳栓型マイクロホン[15]で構成した。測定信号は Swept-Sine Signal (2^{18} サンプル) である。測定信号および録音のサンプリング周波数は 48 kHz である。測定時, 被験者に対して椅子に座り背筋を伸ばして正面を向いて静止するように指示した。被験者の頭部は固定していないが頭部および胴体に傾きがないかについて, 被験者の前方および側方から 2 名の実験者が目測で確認した。被験者は 20 代の 29 名 (男性 27 名, 女性 6 名) である。

3.4 頭部伝達関数の第 1 第 2 ノッチの抽出

3.3 節で HRTF の測定を行った 28 名の 58 耳について, 正面方向の N1,N2,P1 を抽出した。N1, N2, P1 は耳介で発生しているため[19-21], hrir の初期部分から以下のアルゴリズムにより抽出した。

- 1) hrir の振幅の絶対値が最大となるサンプルを検出する。
- 2) hrir を 4 次 96 ポイントの blackman-harris 窓で切り取る。ただし, 1)で抽出した最大サンプルを時間窓の中心に合わせる。
- 3) 全ての要素を 0 とした 512 ポイントの配列を用意し, 2)で切り取った hrir を上書きする。ただし, 最大振幅となるサンプルを配列の 257 番目に合わせる。
- 4) FFT により 512 ポイントの配列の振幅スペクトルを求める。さらに, 差分法により極大値および極小値を検出する。
- 5) 3000 Hz 以上の帯域で最も周波数の低い極大値を P1 とし, それより高い帯域で最も周波数の低い 2 つの極小値を N1, N2 とする。

結果として, 54 耳 (26 人の両耳と 2 人の片耳) の N1, N2, P1 を抽出した。残りの 2 耳については, スペクトラルノッチが 1 つしか抽出できなかった。N1, N2,P1 周波数の分布範囲はそれぞれ, 5719–9563 Hz (0.74 oct.), 8250–13500 Hz (0.71 oct.), 3469–4313 Hz (0.31 oct.)であった。これらの結果は, P1 周波数の個人差は N1, N2 よりも小さく,

N1 と N2 の分布範囲は互いに重なっていることを示している。

上昇角知覚における正面方向の P1 周波数の丁度可知差(JND: just-noticeable difference) は 0.35 oct. (上側周波数) および 0.47 oct. (下側周波数) である[72]。したがって、P1 周波数の個人差は上昇角知覚において影響しないとみなせる。そこで、以降 P1 は議論の対象から外すことにする。一方、N1,N2 の JND は 0.1–0.2 oct. であり[51]、これらの個人差(0.74, 0.71 oct.)は上昇角知覚において十分影響を及ぼすと考えられる。

3.5 耳介形状の計測

Takemoto *et al.*[20]は 4 人の被験者の頭部形状から FDTD 法により HRTF を算出し、4 人の N1, N2 周波数はそれぞれ異なるが、ピーク・ノッチの基本パターンは耳介により生じており、N1 周波数では耳介の腔に 1 つもしくは 2 つの腹と 1 つの節が出現していると報告した。これらの知見は、N1,N2 周波数の個人差は受聴者の耳介形状に起因することを示唆している。

そこで、本研究では Algazi *et al.* [61]と同様の 10 種類の耳介形状パラメータを採り上げ (図 3.1), その個人差を分析した。ただし、彼らを用いたパラメータの 1 つである q (pinna flare angle)は実際には計測が困難であるので、その代わりとして x_4 (width of the helix)を用いた。

54 耳の耳型より 9 種類の耳介パラメータをノギスで計測した。耳介の傾き(x_a)は被験者の横顔の写真から計測した。

計測結果を表 3.1 に示す。各パラメータの分布の幅は 10–25 mm であった。また、耳介の傾き(x_a)は 4–40°であった。

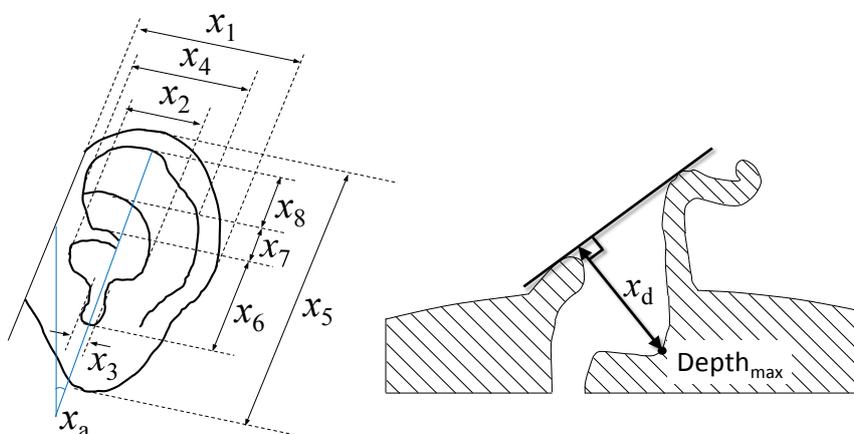


図 3.1 耳介形状の計測箇所

表 3.1 耳介形状の計測結果 [mm]

Pinna	Width of				Length of				Depth of concha	Tilt of pinna(°)
	pinna	concha	incisura intertragica	helix	pinna	concha	cymba conchae	scapha		
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_d	x_a
1	37.0	20.1	8.3	29.7	69.1	21.0	9.3	16.8	15.2	18
2	35.7	19.6	6.5	29.4	72.1	21.5	9.8	15.5	14.9	24
3	33.0	19.7	5.9	27.0	66.3	21.8	6.4	19.8	15.8	29
4	37.5	15.8	5.5	24.7	75.8	19.3	9.9	18.8	13.8	30
5	35.9	18.6	9.0	24.7	73.8	22.9	6.2	19.6	11.9	16
6	34.2	19.5	9.8	26.3	72.0	22.3	9.3	17.9	12.1	17
7	34.7	14.9	5.3	20.1	71.3	21.4	5.9	20.8	16.7	7
8	31.9	16.8	6.8	25.2	70.1	20.2	4.5	21.4	17.2	12
9	34.4	17.8	7.0	25.7	71.4	22.3	5.3	19.4	13.3	17
10	35.0	18.8	9.0	26.1	63.4	22.2	4.7	20.8	14.1	18
11	34.4	18.2	8.3	24.8	64.4	20.4	7.1	20.8	15.0	19
12	38.0	17.6	8.3	28.3	67.0	17.7	8.8	19.2	12.5	23
13	35.1	19.1	8.4	26.3	68.5	17.1	8.1	19.8	13.8	7
14	34.8	18.3	8.7	22.2	66.4	18.8	6.6	22.5	14.1	23
15	35.2	18.7	8.3	27.1	67.0	18.6	8.3	22.0	13.8	24
16	36.1	16.5	5.5	27.3	64.3	19.3	5.5	17.3	14.0	30
17	37.5	16.6	5.5	29.7	66.4	18.7	8.0	17.2	13.8	40
18	36.6	21.2	8.4	27.0	67.2	19.3	7.5	18.4	11.3	27
19	36.1	19.8	7.5	28.1	66.4	19.2	8.5	18.3	11.1	21
20	35.7	15.6	5.8	25.7	69.2	20.5	8.6	20.3	13.7	28
21	35.8	15.4	6.4	26.8	70.8	20.9	9.4	21.2	13.1	36
22	36.1	20.5	6.7	27.0	64.0	21.0	6.0	20.2	14.8	32
23	33.4	16.5	5.9	27.5	63.8	22.5	4.3	21.1	13.8	27
24	43.8	20.2	8.3	31.6	78.3	21.9	8.0	24.1	14.1	27
25	42.1	19.2	7.3	31.8	77.8	19.0	9.3	23.1	14.4	20
26	36.9	14.8	6.7	27.5	72.8	24.1	7.5	19.4	13.0	23
27	36.1	17.4	7.5	24.3	70.4	22.9	9.3	17.4	13.8	31
28	31.7	16.6	8.9	19.0	69.2	23.3	7.7	15.9	16.0	25
29	32.8	16.7	8.7	18.1	69.3	20.4	9.1	17.1	15.7	24
30	36.0	20.1	10.8	28.3	64.0	22.0	5.8	16.7	13.8	27
31	34.0	21.8	11.9	26.9	64.4	22.0	10.3	14.2	13.1	36
32	40.6	19.1	10.2	28.0	75.8	25.1	2.6	21.1	14.1	23
33	36.4	18.4	7.6	26.2	83.2	24.9	4.7	22.1	13.8	29
34	31.2	19.3	10.9	21.6	63.5	20.4	5.9	15.4	13.3	13
35	32.5	19.9	10.1	23.0	65.1	21.3	8.2	15.9	11.7	14
36	33.1	21.0	9.0	23.1	58.2	18.4	3.5	14.0	9.9	32
37	32.7	20.4	7.2	26.6	59.1	17.8	5.1	16.9	9.7	16
38	35.7	19.7	7.4	29.0	66.9	20.2	5.3	17.6	11.8	19
39	35.0	21.8	7.3	28.8	68.9	19.5	7.7	20.0	12.5	28
40	34.5	17.5	8.4	24.3	66.1	23.0	5.9	16.5	14.9	29
41	34.5	19.9	8.1	25.3	68.2	23.1	6.3	17.9	15.0	17
42	35.2	19.0	10.0	26.6	69.9	21.6	4.0	20.3	12.6	18
43	34.5	19.1	6.1	28.3	72.5	21.0	6.0	20.7	15.2	20
44	37.4	20.9	7.5	29.4	73.8	19.4	9.7	23.2	13.6	33
45	35.3	20.0	11.2	25.2	63.4	24.1	3.5	18.5	15.6	14
46	33.2	18.5	9.9	22.8	68.7	24.5	3.9	16.5	17.6	10
47	32.9	18.5	10.4	24.0	62.4	21.2	5.8	18.9	14.8	4
48	35.9	19.2	9.0	26.1	63.2	20.1	7.1	17.1	14.0	17
49	32.9	20.0	9.3	23.2	64.5	23.5	2.9	13.2	17.6	14
50	37.3	18.1	8.3	23.9	65.7	21.8	4.7	14.9	17.4	15
51	35.7	18.1	8.8	26.2	66.7	22.3	6.1	18.8	13.4	20
52	38.4	17.9	10.1	26.4	65.4	20.4	5.6	20.8	11.3	9
53	36.9	19.9	11.5	28.8	66.8	21.0	6.5	19.4	13.6	39
54	38.0	21.7	8.1	28.2	68.8	22.8	7.5	21.4	14.8	28
Min	31.2	14.8	5.3	18.1	58.2	17.1	2.6	13.2	9.7	4
Max	43.8	21.8	11.9	31.8	83.2	25.1	10.3	24.1	17.6	40

3.6 耳介形状による第1第2ノッチ周波数の推定

3.6.1 重回帰分析

上記の54耳を用い、10種類の耳介形状パラメータを説明変数、正面方向のN1,N2周波数を目的変数として重回帰分析を行った(式3.1)。

$$f(s)_{N1,N2} = a_1x_1 + a_2x_2 + \dots + a_nx_n + b \quad [\text{Hz}] \quad (3.1)$$

ここで、 S, a_i, b, x_i はそれぞれ被験者、回帰係数、定数、耳介形状パラメータを表す。

なお、多重共線性の有無については、VIF(Variance Inflation Factor)によって確認した。本研究で用いたすべての説明変数のVIF(10種類)の値は10未満であったため説明変数間には多重共線性はないとみなせる。

3.6.2 重回帰モデルの精度

重回帰分析の結果、N1およびN2の重相関係数は0.84および0.87であった。しかし、各説明変数のp値については、N1では x_1, x_4, x_7 で0.05を超え、N2では x_1, x_2, x_3, x_4, x_a で0.05を超えた。これらのパラメータはN1,N2に寄与しているとは考えにくい。そこで、説明変数の個数を変えながら全ての説明変数の組み合わせについて重回帰分析を行い、全ての説明変数のp値が0.05未満かつ重相関係数が最大となる組み合わせを求めた。その結果、N1については6種類($x_2, x_3, x_6, x_8, x_d, x_a$)、N2については3種類(x_6, x_8, x_d)の耳介形状パラメータが選ばれた。つまり、N1においては耳介の窪みの幅、長さ、深さと耳介の傾きが寄与し、N2においては耳介の窪みの長さ、深さが寄与していると言える。重回帰係数、p値、95%信頼区間を表3.2に示す。

重回帰モデルの統計量を表3.3に示す。N1,N2の重相関係数はそれぞれ0.81、0.82であった。平均絶対残差はそれぞれ0.07および0.08 oct.であった。絶対残差がJND以内となる確率はN1,N2共通に91%であった。JNDは先述の通り0.1-0.2 oct.と考えられるので0.15 oct.とした。

以上の結果は、ここで提案した重回帰モデルがほぼJND以内の誤差でN1およびN2を推定できることを示している。

表 3.2 重回帰係数, p 値, 95%信頼区間

	Regression coefficient		P-value		95% confidence intervals			
					lower	upper	lower	upper
	N1	N2	N1	N2	N1	N2		
a_1								
a_2	116.9		1.6.E-02		22.9	210.9		
a_3	-157.5		4.7.E-03		-264.2	-50.8		
a_4								
a_5								
a_6	-183.4	-327.0	8.3.E-05	2.9.E-07	-269.1	-97.8	-438.0	-216.0
a_7								
a_8	-93.2	-245.0	2.3.E-03	4.4.E-08	-151.5	-34.9	-321.3	-168.6
a_d	-131.4	-172.8	4.0.E-03	3.7.E-03	-218.7	-44.2	-286.9	-58.7
a_a	-48.7		7.2.E-07		-65.8	-31.6		
b	14906.4	23903.1	9.2.E-14	2.0.E-22	12019.9	17792.9	21079.9	26726.3

表 3.3 重回帰モデルの統計値

	Correlation coefficient	Significance level	Absolute mean residual error		Probability that residual error less than 0.15 octaves [%]
			[Hz]	[oct.]	
N1	0.81	1.1.E-09	357	0.07	91
N2	0.82	5.4.E-12	550	0.08	91

3.7 ナイーブな被験者の第 1, 第 2 ノッチ周波数の推定精度

3.7.1 物理的精度

a) 正面方向の第 1 第 2 ノッチ周波数の推定

重回帰分析に含まれていない被験者の N1,N2 周波数を推定し,その精度を検証した.

被験者は男性 3 人(OIS, TCY, MTZ)および女性 1 人(CKT)であり,年齢は 21-25 歳である. N1,N2 周波数の推定に必要な 6 箇所の耳介形状パラメータを実耳から計測した. その結果を表 3.4 に示す. これらと重回帰式を用いて正面方向の N1,N2 周波数を推定した. また, 実測 HRIR から N1,N2 を抽出し, その周波数を求めた. 推定値, 実測値, および推定誤差を表 3.5 に示す. N1,N2 の推定誤差は 8 耳 (4 名の両耳) すべてにおいて JND 未満であった. 比較的大きな誤差は被験者 CKT の左耳(0.10 oct.), MTZ の左耳(0.09 oct.)および TCY の右耳(0.09 oct.)で観測された.

表 3.4 4名のナイーブな被験者の耳介形状

Subject	Ear	Width of		Length of		Depth of	Tilt of
		concha	incisura intertragica	concha	scapha		
		x_2	x_3	x_6	x_8	x_d	x_a
OIS	L	19.2	7.9	23.3	19.5	12.9	28
	R	17.0	8.7	21.0	20.1	13.5	18
TCY	L	14.4	8.0	23.0	20.3	12.5	29
	R	14.3	7.6	22.8	20.4	12.9	23
CKT	L	15.7	7.4	19.7	18.3	13.9	22
	R	15.0	8.2	19.6	18.3	12.1	21
MTZ	L	17.6	7.6	22.6	17.2	11.2	28
	R	18.1	7.2	21.5	18.0	13.9	24

表 3.5 4名の被験者の N1,N2 周波数の推定結果

Subject	Ear	Estimated frequency[Hz]		Extracted frequency[Hz]		Residual error[oct.]	
		N1	N2	N1	N2	N1	N2
OIS	L	6749	9273	6938	9375	-0.04	-0.02
	R	7147	9779	6938	9281	0.04	0.08
TCY	L	6163	9249	6094	9656	0.02	-0.06
	R	6481	9221	6094	9375	0.09	-0.02
CKT	L	7358	10576	6844	10406	0.10	0.02
	R	7454	10920	7219	11250	0.05	-0.04
MTZ	L	7182	10364	6750	10875	0.09	-0.07
	R	7271	10061	7313	10125	-0.01	-0.01

b) best-matching HRTF の選出

次に、推定した N1, N2 周波数に基づいてデータベースから適合する HRTF を選出する方法を提案し、音像定位実験によりその有効性を検証した。ここで、推定した N1, N2 周波数に最も近い周波数をもつ HRTF を best-matching HRTF と呼ぶ。

HRTF データベースから best-matching HRTF を選出する物理指標として、NFD(notch frequency distance)を用いた[73]。NFD は被験者 j の $HRTF_j$ と被験者 k の $HRTF_k$ の距離を octave スケールにより以下の式で表したものである。

$$\text{NFD}_1 = \log_2 \left\{ \frac{f_{N1}(\text{HRTF}_j)}{f_{N1}(\text{HRTF}_k)} \right\} \text{ [oct.]} \quad (3.2)$$

$$\text{NFD}_2 = \log_2 \left\{ \frac{f_{N2}(\text{HRTF}_j)}{f_{N2}(\text{HRTF}_k)} \right\} \text{ [oct.]} \quad (3.3)$$

$$\text{NFD} = |\text{NFD}_1| + |\text{NFD}_2| \text{ [oct.]} \quad (3.4)$$

ここで、 f_{N1} および f_{N2} はそれぞれ $N1$ および $N2$ 周波数を表す。

正面方向の $N1, N2$ 周波数の推定値との NFD が最小となる HRTF をデータベースから選出し、**best-matching HRTF** とした。ただし、左耳と右耳では個別に選出した。加えて、正面以外の上半球正中面の方向の **best-matching HRTF** として、正面方向の **best-matching HRTF** の提供者の HRTF を用いた。

今回用いたデータベースは日本人成人 120 耳の上半球正中面 7 方向の HRTF で構成されている。120 耳のうち 54 耳は重回帰分析に用いたものである。なお、このデータベースは著者らの研究室の web サイトで公開している[48]。

4 人の被験者(OIS, TCY, CKT, MTZ)の正面方向の **best-matching HRTF** を $N1, N2$ 周波数の推定値より選出した。図 3.2 に 4 人の被験者本人の HRTF と **best-matching HRTF** の振幅スペクトルを示す。**best-matching HRTF** (点線) の $N1, N2$ 周波数 (●, ▲) と被験者本人の HRTF (実践)の $N1, N2$ 周波数 (○, △) は近接している。また、ほとんど全ての耳において、**best-matching HRTF** (点線) と被験者本人の HRTF (実線) で同様のスペクトル構造が観察された。しかし、被験者 CKT の左耳の **best-matching HRTF** は本人の HRTF と類似しているとは言えない。**best-matching HRTF** の $N1$ は本人のそれと比較して浅く、 $N2$ は深い。提案法ではノッチのレベルを考慮していないので、このようなことは起こり得る。

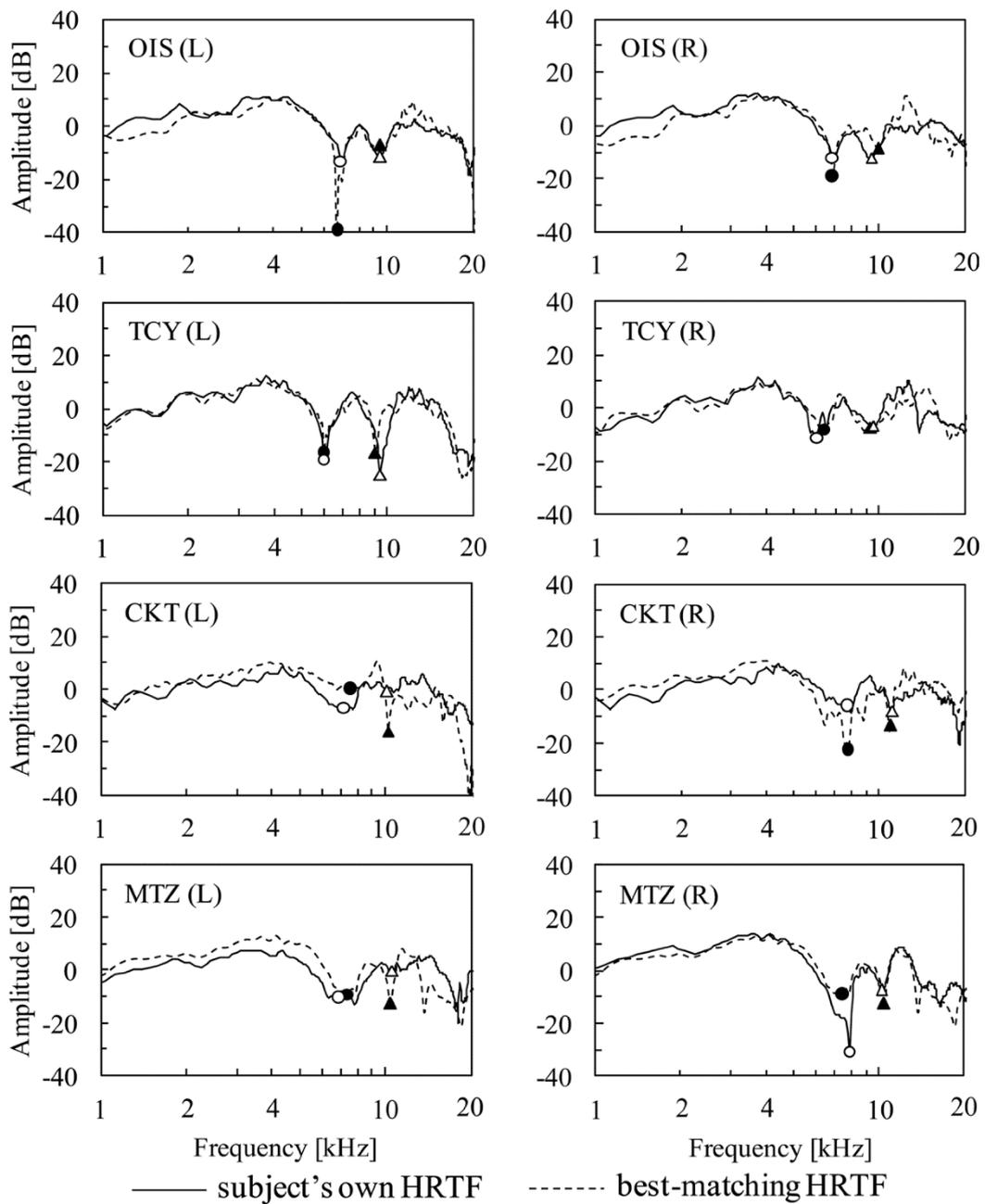


図 3.2 被験者本人の HRTF と best-matching HRTF の振幅スペクトル

● : N1(best-matchingHRTF), ○N1 (本人 HRTF),

▲ : N2(best-matchingHRTF), △N2 (本人 HRTF)

表 3.6 に、正面方向の被験者本人の HRTF および best-matching HRTF の N1,N2 周波数を示す。残差は、耳介形状による N1,N2 周波数の推定誤差とデータベースから best-matching HRTF を選出する際の誤差の重ね合わせである。N1,N2 の残差は 8 耳全てにおいて JND 未満であった。比較的大きな誤差は被験者 CKT の左耳(0.11 oct.), MTZ の左耳(0.10 oct.)および TCY の右耳(0.09 oct.)で観察された。これは、表 3.5 と同じ傾向であり、推定誤差が支配的であると言える。

表 3.6 正面方向における本人 HRTF と best-matching HRTF の N1,N2 周波数

Subject	Ear	Best-matching frequency [Hz]		Extracted frequency [Hz]		Residual error [oct.]	
		N1	N2	N1	N2	N1	N2
OIS	L	6844	9375	6938	9375	-0.02	0.00
	R	6938	9844	6938	9281	0.00	0.08
TCY	L	6094	9188	6094	9656	0.00	-0.07
	R	6469	9188	6094	9375	0.09	-0.03
CKT	L	7406	10594	6844	10406	0.11	0.03
	R	7500	10875	7219	11250	0.06	-0.05
MTZ	L	7219	10313	6750	10875	0.10	-0.08
	R	7219	10313	7313	10125	-0.02	0.03

図 3.3 は、被験者本人の HRTF と best-matching HRTF の N1-N2 平面での散布図である。120 耳の布置は広がっているが、いずれの被験者においても本人の HRTF と best-matching HRTF は近接している。ただし、被験者 CKT と MTZ の左耳は他の被験者と比較して距離が大きい。

次に、上半球正中面 7 方向における残差を示す (表 3.7)。表 3.6 で示したように 0°での残差は JND 未満であった。しかし、N1 については、被験者 OIS の左耳(0.20 oct.), 150°の右耳(0.16 oct.), 被験者 CKT の 120°の右耳(-0.16 oct.)において残差は JND を上回った。

以上より、正面方向だけでなく、他の上半球正中面のほとんどの方向において、best-matching HRTF は被験者本人の HRTF と類似のスペクトルを有していると考えられる。

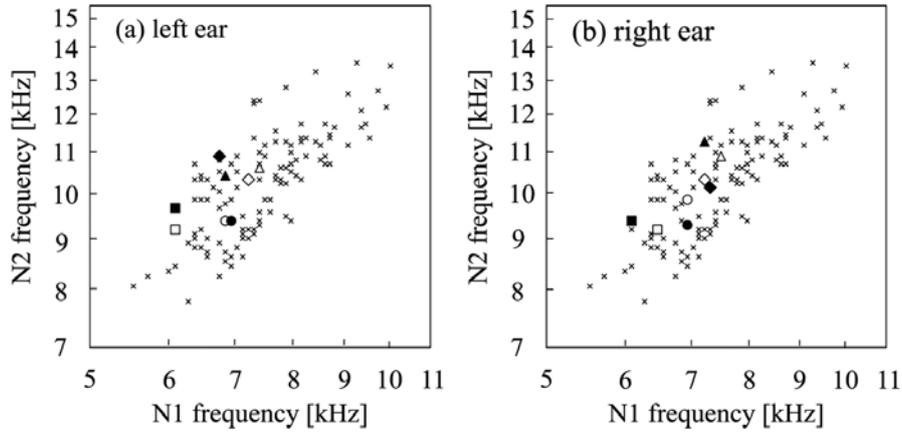


図 3.3 正面方向の本人 HRTF と best-matching HRTF の N1,N2 周波数の散布図

(a)左耳, (b)右耳. ● : OIS(best-matching), ○ : OIS(own), ■ : TCY(best-matching), □ : TCY(own) , ▲ : CKT(best-matching), △ : CKT(own), ◆ : MTZ(best-matching), ◇ : MTZ(own), × : other HRTFs in the database.

表 3.7 正中面 7 方向における被験者本人の HRTF と best-matching HRTF の N1, N2 周波数の差 [oct.]

Subject	Ear	Notch	Target angle (°)						
			0	30	60	90	120	150	180
OIS	L	N1	-0.02	0.10	-0.09	-0.03	-0.04	0.13	0.20
		N2	0.00	0.14	0.07	-0.04	0.06	0.11	0.08
	R	N1	-0.02	0.05	0.00	0.08	0.06	0.16	0.00
		N2	0.08	0.06	0.01	0.03	0.04	0.00	0.00
TCY	L	N1	-0.02	-0.04	-0.02	0.00	-0.07	0.01	-0.05
		N2	-0.09	0.01	-0.03	0.06	0.02	-0.02	0.08
	R	N1	0.09	0.00	0.00	-0.07	-0.05	-0.01	-0.03
		N2	-0.03	0.08	0.00	0.08	0.00	-0.10	0.05
CKT	L	N1	0.11	0.13	-0.03	-0.08	-0.09	-0.07	-0.09
		N2	0.03	-0.01	0.06	0.11	0.04	-0.07	-0.01
	R	N1	0.06	0.02	-0.12	-0.04	-0.16	-0.03	-0.06
		N2	-0.05	-0.08	-0.14	-0.03	-0.04	0.12	0.04
MTZ	L	N1	0.10	0.07	0.08	-0.06	-0.02	0.07	0.08
		N2	-0.08	0.08	-0.03	-0.07	-0.06	0.10	0.06
	R	N1	-0.02	0.09	0.00	-0.06	0.05	-0.09	-0.08
		N2	0.03	0.03	0.01	-0.01	-0.03	0.05	-0.05

3.7.2 心理的評価

A 実験方法

音像定位実験は防音室で行った。室の内寸は 4.6 m (W) × 5.8 m (D) × 2.8 m (H) で、暗騒音は 19.5 dB(A) であった。実験システムは、ノート PC (DELL XPS M1330)、オーディオインターフェース (RME Fireface 400)、アンプ (Marantz PM4001)、オープンタイプヘッドホン (AKG K1000)、耳栓型マイクロホン [15]、A/D コンバータ (Roland M-10MX) で構成した。

Møller *et al.* [26] は、外耳道入口を開放した状態における鼓膜での音圧は、閉塞した状態の外耳道入口の音圧に以下の補正 G を施すことによって得られることを示した。

$$G = \frac{1}{M \cdot PTF} \cdot \frac{Z_{\text{ear canal}} + Z_{\text{headphone}}}{Z_{\text{ear canal}} + Z_{\text{radiation}}} = \frac{1}{M \cdot PTF} \cdot PDR \quad (3.5)$$

ここで、 M は耳栓型マイクロホンの伝達関数、 PTF は閉塞した外耳道入口で測定したヘッドホンの伝達関数、 $Z_{\text{ear canal}}$ と $Z_{\text{headphone}}$ はそれぞれ外耳道およびヘッドホンのインピーダンス、 $Z_{\text{radiation}}$ は外耳道から外側をみた放射インピーダンスである。式中の PDR は *pressure division ratio* の略であり、彼らは本実験で用いたヘッドホン (AKG K-1000) の PDR は 1 とみなせると報告している。

この原理に則り刺激を提示した。まず、マイクロホンの振動板が外耳道入口に位置するように被験者に耳栓型マイクロホンを装着した。次に、被験者にヘッドホンを装着して 12 次の M 系列信号 (サンプリング周波数: 48 kHz) を提示し、ヘッドホンから耳栓型マイクロホンまでの伝達関数 $M \cdot PTF$ を測定した。さらに、ヘッドホンを装着したまま耳栓型マイクロホンだけを外し、以下の式で作成した刺激 $P_{l,r}(\omega)$ をヘッドホンから提示した。

$$P_{l,r}(\omega) = S(\omega) \cdot \frac{HRTF_{l,r}(\omega)}{M_{l,r}(\omega) \cdot PTF_{l,r}(\omega)} \quad (3.6)$$

ここで、 $S(\omega)$ 、 l, r はそれぞれ音源信号、左耳、右耳を表す。

音源信号は 200 Hz–17 kHz の広帯域ホワイトノイズである。HRTF には、被験者本人の実測 HRTF と *best-matching HRTF* を用いた。

目標方向は上半球正中面における 30° 間隔の 7 方向である。刺激の提示音圧レベルは、被験者の左右の外耳道入口で 63 dB とした。両耳間時間差およびレベル差はゼロに設定した。刺激の提示時間は 1.2 秒 (立上がり、立下がりそれぞれ 0.1 秒を含む) である。

被験者がターゲット方向を推測することを避けるため、回答方法にはマッピング法を用いた(図 3.4)。被験者はノート PC の画面上の正中面を表す円周上に、音像方向をマウスでクリックして回答した。頭内定位した場合は、頭内定位を回答するボタンを押した。刺激はランダムな順に提示し、各被験者は各刺激について 10 回ずつ回答した。被験者は前節で best-matching HRTF を求めた OIS, TCY, CKT, MTZ の 4 名である。

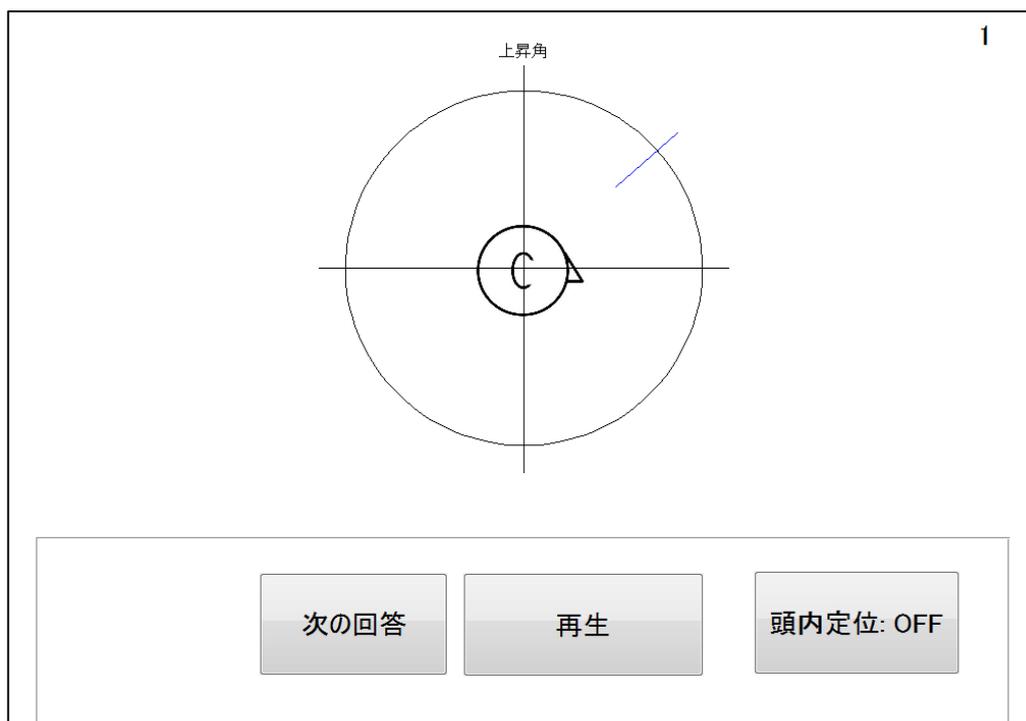


図 3.4 マッピング法による回答画面

B 実音源を用いた音像定位実験

HRTF を用いた音像定位実験に先立って実音源による音像定位実験を行った。この実験の目的は被験者の正中面音像定位能力を確認することである。実験は無響室で行った。音源信号は 200 Hz–17 kHz の広帯域ホワイトノイズであり、上半球正中面の 30° 間隔の 7 方向に設置した小型スピーカ(FOSTEX FE83E)からランダムな順に提示した。スピーカから被験者の頭部中心までの距離は 1.2 m である。なお、中心周波数 250 Hz から 16 kHz までの 1/3 octave バンドのレベルの偏差が 1 dB 以内になるようにイコライザで調整している。無響室は消灯し、回答に必要な手元灯のみ使用した。被験者はマッピング法により解答用紙に音像方向を記入した。各被験者は各刺激に対し 10 回ずつ回答した。

C 実験結果

a) 実音源と各 HRTF に対する回答

図 3.5 に実音源, 被験者本人の HRTF, best-matching HRTF に対する 4 人の被験者の回答を示す. 縦軸は知覚した上昇角, 横軸は目標の上昇角である. 円の直径は回答の頻度に比例している. なお, 図中の描画において, 回答方向は 5° 間隔で丸めている.

被験者 OIS では, 実音源に対する回答(a)は対角線を跨いで S 字カーブ状に分布している. ただし, 目標方向 60° および 120° に対する回答はやや上方にシフトしている. best-matching HRTF (c) に対しては, N1, N2 周波数を推定した 0° 方向では, 本人 HRTF とほぼ同じ方向に知覚している. また, 目標方向 30, 60, 180° での回答の分布は本人 HRTF と同様であった. 一方, 目標方向 90° および 120° では 120–150° に知覚する傾向があり, 目標方向 150° での回答は上昇する傾向があった.

被験者 TCY では, 実音源に対する回答(a)は目標方向 0–60° および 180° では対角線上に布置している. しかし, 目標方向 90° および 120° では後方に, 150° では上方に知覚する傾向がある. 本人の HRTF (b) に対しては目標方向 120° および 150° では回答がばらつくが, 他の方向では対角線上に布置している. best-matching HRTF (c) に対しては, N1, N2 周波数を推定した目標方向 0° では, 本人 HRTF とほぼ同じ方向に知覚している. 30, 60, 180° でも目標方向付近に知覚している. 目標方向 90° では回答のばらつきは本人の HRTF より大きい, 実音源と同程度であった. 120° では回答のばらつきは本人の HRTF と同程度であった. 150° では回答のばらつきは本人 HRTF よりも小さい.

被験者 CKT では, 実音源に対する回答(a)は概ね対角線上に布置している. ただし, 目標方向 0° において一度だけ上方に回答した. 本人の HRTF (b) に対しては, 概ね目標方向に知覚している. ただし, 目標方向 150° では回答のばらつきは実音源よりも大きい. best-matching HRTF (c) に対しては, N1, N2 周波数を推定した目標方向 0° では, 本人 HRTF とほぼ同じ方向に知覚している. 目標方向 150° および 180° でも本人 HRTF と同等の分布であった. 30° では回答は上方にシフトする傾向がある. 他の 3 方向(60, 90, 120°)では, 回答のばらつきは本人 HRTF より大きい.

被験者 MTZ では, 実音源に対する回答(a)は, 対角線を跨いで S 字カーブ状に分布している. 本人の HRTF (b) に対しても実音源同様に S 字カーブを描いている. ただし, 目標方向 90° および 120° に対する回答はやや後方にシフトしている. best-matching HRTF (c) に対しては, N1, N2 周波数を推定した目標方向 0° では, 本人 HRTF とほぼ同じ方向

に知覚している。目標方向 60, 120, 150, 180°においても回答の分布は本人 HRTF と同等であった。ただし、30°および90°では回答のばらつきは本人 HRTF より大きい。

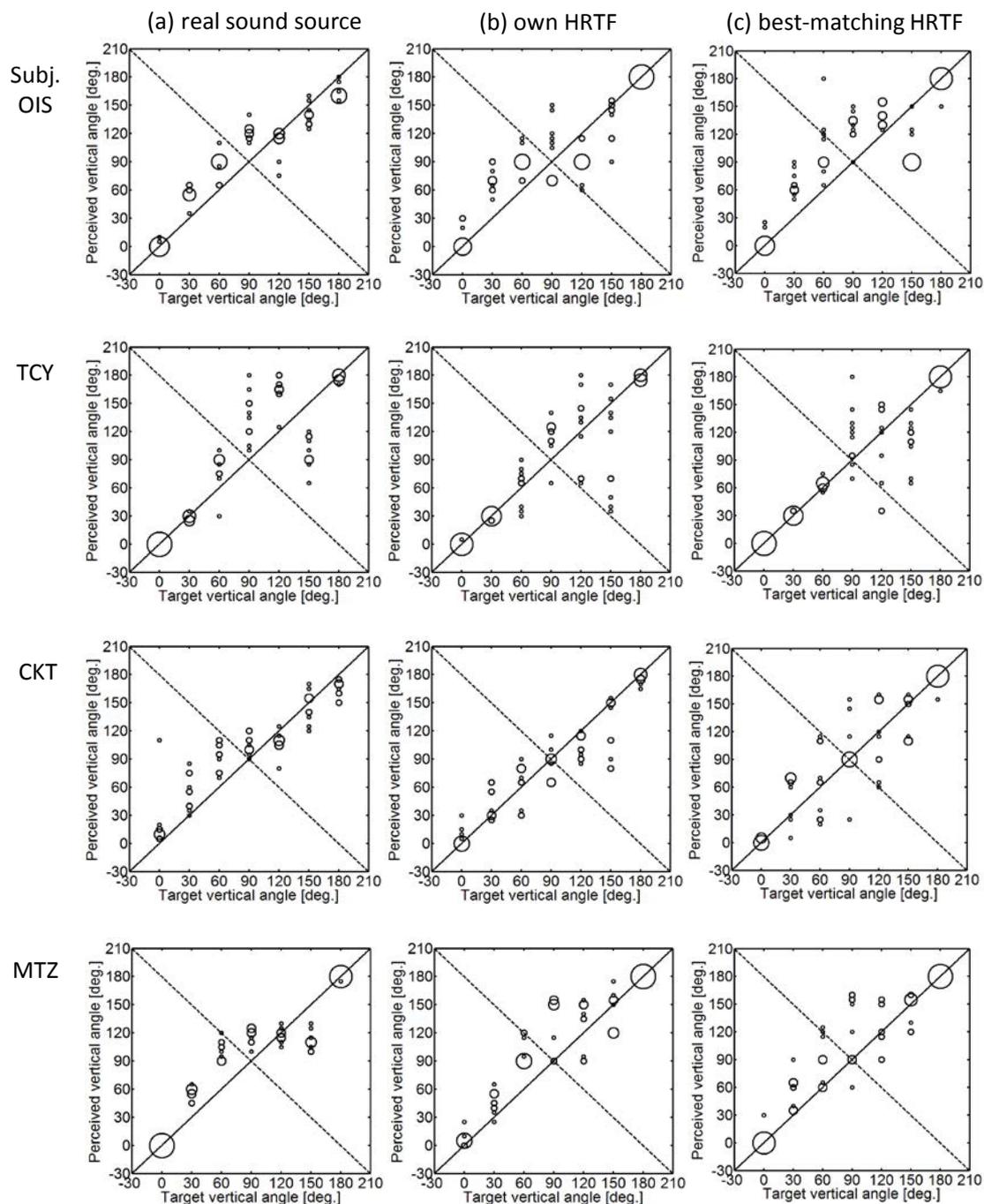


図 3.5 実音源、本人の HRTF、best-matching HRTF による

正中面音像定位実験の結果

b) 平均定位誤差

表 3.8 に平均定位誤差を示す. ここで, 平均定位誤差は目標方向の上昇角と知覚した上昇角の差の絶対値の平均値と定義する.

表 3.8 平均定位誤差[°]

Subject	HRTF	Target vertical angle (°)							Ave.
		0	30	60	90	120	150	180	
OIS	real sound source	2.5	25.9	27.4	32.0	9.5	12.4	17.5	18.2
	own HRTF	8.7	40.4	30.5	28.1	29.7	16.4	0.5	22.0
	best-matching HRTF	5.2	36.7	44.3	39.8	20.4	47.3	3.4	28.2
TCY	real sound source	0.4	2.8	25.7	45.7	44.0	52.2	2.8	24.8
	own HRTF	0.7	1.8	17.3	29.5	34.2	56.8	3.0	20.5
	best-matching HRTF	0.3	1.9	4.9	30.5	36.7	40.0	2.4	16.7
CKT	real sound source	20.8	25.2	32.9	14.8	13.3	13.2	15.4	19.4
	own HRTF	5.9	13.6	19.1	12.1	17.4	29.0	4.1	14.4
	best-matching HRTF	2.1	29.2	30.8	21.3	31.9	17.3	3.5	19.5
MTZ	real sound source	0.9	25.1	41.6	26.9	5.4	39.0	0.8	20.0
	own HRTF	5.9	16.6	39.2	46.2	25.4	17.1	1.1	21.6
	best-matching HRTF	3.3	24.8	27.9	38.5	20.6	13.1	0.5	18.4

被験者 OIS では, best-matching HRTF の目標方向 0°および 180°の平均定位誤差は 5.2°および 3.4°であった. これらの値は本人の HRTF と同程度である. しかし, 150°での平均定位誤差は 47.3°であり, 本人 HRTF の約 3 倍となった. これは表 3.7 に示した本人 HRTF と best-matching HRTF の N1 周波数の比較的大きな差 (右耳: 0.16 oct., 左耳: 0.13 oct.) に依るものと解釈できる. 表 3.7 にはもう 1 つの大きい N1 周波数の差異 (180°における左耳: 0.20 oct.) が示されているが, この場合は右耳での差異が N1, N2 とともに 0.00 oct. であり, これが小さな平均定位誤差(3.4°)に貢献していると考えられる.

被験者 TCY では, best-matching HRTF の目標方向 0°および 180°の平均定位誤差は 0.3°および 2.4°であった. これらの値は本人の HRTF と同等である. また, 他の 5 方向においても best-matching HRTF の平均定位誤差は本人 HRTF のそれと同等であった.

被験者 CKT では, best-matching HRTF の目標方向 0°および 180°の平均定位誤差は 2.1°および 3.5°であった. これらの値は本人 HRTF と同程度である. しかし, 30°での平均定位誤差は 29.2°であり, 本人 HRTF の約 2 倍となった. これは表 3.7 に示した本人 HRTF と best-matching HRTF の N1 周波数の比較的大きな差異 (左耳: 0.13 oct.) に依るものと解釈できる. また, 120°での平均定位誤差は 31.9°であり, 本人 HRTF の約 2 倍となっ

た. これも本人 HRTF と best-matching HRTF の N1 周波数の比較的大きな差異 (右耳 : -0.16 oct.) に依るものと解釈できる.

被験者 MTZ では, best-matching HRTF の目標方向 0°および 180°の平均定位誤差は 3.3° および 0.5°であった. これらの値は本人 HRTF と同程度である. また, 他の 5 方向においても best-matching HRTF の平均定位誤差は本人 HRTF のそれと同等であった.

c) 前後誤判定率

表 3.9 に前後誤判定率を示す. ここで, 前後誤判定率は上半球正中面において目標方向と異なる四分円に音像を知覚した率と定義する.

全ての被験者において, best-matching HRTF の目標方向 0°および 180°の前後誤判定率は本人 HRTF 同様 0%であった. 他の 5 方向についても best-matching HRTF の前後誤判定率は本人 HRTF と同等であった. ただし, 被験者 OIS の 150°, CKT の 60°, MTZ の 120°では best-matching HRTF の前後誤判定率は本人 HRTF よりも高い.

実音源, 本人 HRTF, best-matching HRTF の 7 方向の平均前後誤判定率について, カイ二乗検定を行った. 表 3.10 に p 値を示す. 全ての p 値は 0.05 よりも大きい. つまり, 実音源, 本人 HRTF, best-matching HRTF の間の前後誤判定率に統計的有意差は認められない.

表 3.9 前後誤判定率

Subject	HRTF	Target vertical angle (°)							Ave.
		0	30	60	90	120	150	180	
OIS	real sound source	0	0	40	-	10	0	0	8.3
	own HRTF	0	20	70	-	20	10	0	20.0
	best-matching HRTF	0	10	60	-	0	30	0	16.7
TCY	real sound source	0	0	10	-	0	20	0	5.0
	own HRTF	0	0	10	-	30	50	0	15.0
	best-matching HRTF	0	0	0	-	30	20	0	8.3
CKT	real sound source	10	0	60	-	10	0	0	13.3
	own HRTF	0	0	10	-	20	30	0	10.0
	best-matching HRTF	0	0	30	-	20	0	0	8.3
MTZ	real sound source	0	0	100	-	0	0	0	16.7
	own HRTF	0	0	100	-	0	0	0	16.7
	best-matching HRTF	0	10	50	-	20	0	0	13.3

表 3.10 カイ二乗検定の結果

Comparison between	Subject	p-value
own HRTF and best-matching HRTF	OIS	0.64
	TCY	0.26
	CKT	0.75
	MTZ	0.61
real sound source and best-matching HRTF	OIS	0.17
	TCY	0.71
	CKT	0.38
	MTZ	0.61
real sound source and own HRTF	OIS	0.067
	TCY	0.068
	CKT	0.57
	MTZ	1.0

d) 頭内定位率

全ての被験者において、best-matching HRTF および本人 HRTF では頭内定位は生じなかった。

e) 音像定位実験結果のまとめ

上記の結果は、受聴者の N1,N2 周波数を推定した正面方向(0°)において best-matching HRTF は本人 HRTF と同等の音像定位精度を有していることを示している。また、best-matching HRTF は後方(180°)においても前方と同様の音像定位精度を有していることを示している。他の5方向については、被験者の一部で本人 HRTF よりも音像定位精度は低下した。

D best-matching HRTF と非個人化 HRTF の定位精度の比較

提案方法の有効性を検証するため、上半球正中面内を目標方向とした定位実験を行った。被験者 OIS と TCY の HRTF を非個人化 HRTF として用いた。実験方法は前節と同様である。図 3.6, 3.7 に OIS と TCY の HRTF に対する各被験者の実験結果を示す。

OIS の HRTF に対する各被験者の分布をみる(図 3.6)。被験者 TCY の回答は、目標方向 0, 180°では、目標方向に定位した。しかし、目標方向 30, 60, 90, 150°では、ばらつきは best-matching HRTF よりも大きい。被験者 CKT の回答は、目標方向 0°では、0, 120, 150°に定位したが、best-matching HRTF では、0°付近に定位していた。目標方向 60, 90,

120, 150°では, 60°から 150°の間に分布した. 目標方向 180°では, best-matching HRTF と同様に目標方向に定位した. 被験者 MTZ では, 目標方向 0, 180°では, 前方と後方に分布が, best-matching HRTF では概ね目標方向に定位していた. 目標方向 30°では, 45–150°にばらついて定位した. 目標方向 60, 90, 120°では, 90–180 にばらついて定位した.

次に, 被験者 TCY の HRTF に対する各被験者の回答をみる (図 3.7). 被験者 OIS では, best-matching HRTF と同様に概ね対角線上に回答が分布した. しかし, 目標方向 0°において前後誤判定が 1 回生じた. 被験者 CKT では, 目標方向 0°において後方に定位した. 目標方向 60–150°では, 60–90°に定位した. 目標方向 180°では, best-matchig HRTF と同様に, 目標方向に定位した. 被験者 MTZ では, 目標方向 0°において後方に定位した. 目標方向 90–150°では, 90–180°に分布した. 目標方向 180°では, 0–180°の広い範囲に分布した.

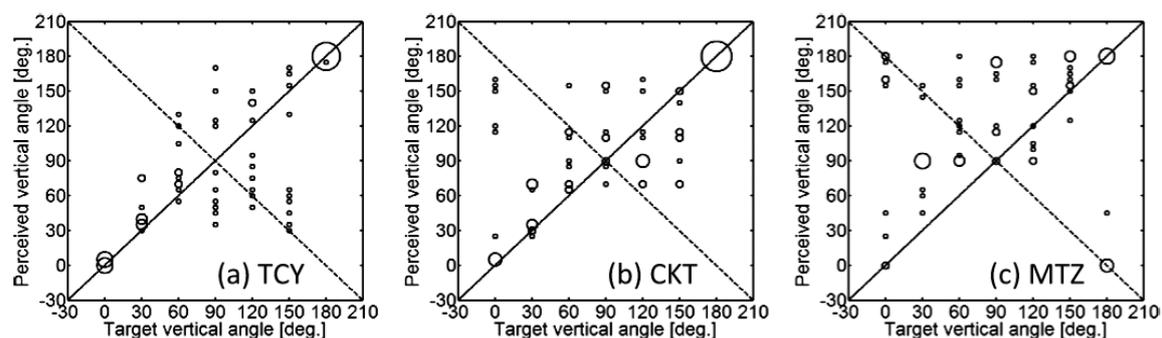


図 3.6 OIS の HRTF に対する各被験者の回答分布

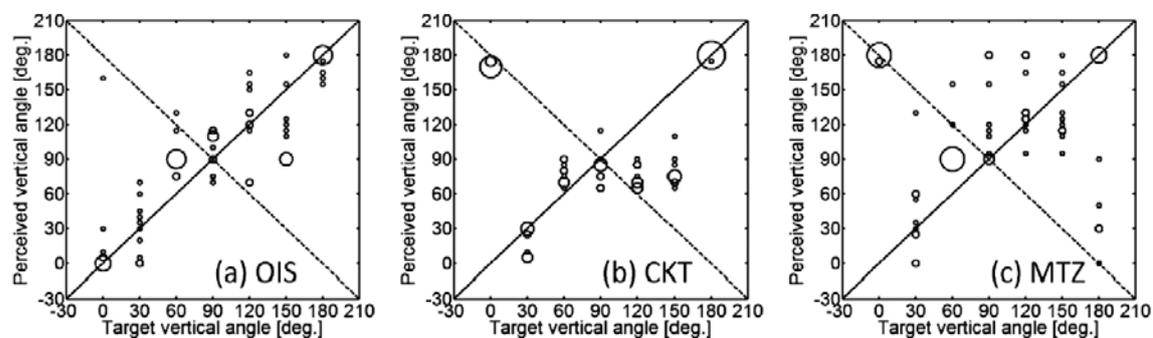


図 3.7 TCY の HRTF に対する各被験者の回答分布

表 3.11 に平均定位誤差, 表 3.12 に前後誤判定率を示す. 被験者 OIS については, TCY の HRTF の平均定位誤差は, 目標方向 0°では best-matching HRTF よりも大きい. これは前後誤判定によるものである. 120, 180°では best-matching HRTF と同等である. 30, 60, 90, 150°では, best-matching HRTF よりも小さい.

被験者 TCY については, OIS の HRTF に対しては, 30, 60, 90, 150°では, best-matching HRTF よりも大きい. その他の目標方向では, best-matching HRTF と同等である.

被験者 CKT については, OIS の HRTF に対しては, 0, 90, 150°では best-matching HRTF よりも大きい. 0°は高い頻度(50%)で生じた前後誤判定によるものである. 30°では, best-matching HRTF よりも小さい. TCY の HRTF に対しては, 0, 120, 150°では, best-matching HRTF よりも大きい. これは高い頻度 (いずれも 100%) で生じた前後誤判定によるものである. 30, 60, 90°では, best-matching HRTF よりも小さい.

被験者 MTZ については, TCY の HRTF の 90°を除き, best-matching HRTF よりも大きい.

非個人化 HRTF の 7 方向の平均定位誤差の平均値は, 被験者 OIS を除き, best-matching HRTF のそれよりも大きい. さらに, 非個人化 HRTF の 7 方向の前後誤判定率の平均値は, 全ての被験者において best-matching HRTF よりも大きい.

各 HRTF に対する前後誤判定率についてカイ二乗検定を行った (表 3.13). その結果, 被験者 OIS の TCY の HRTF と best-matching HRTF の比較を除き, best-matching HRTF の前後誤判定率は他の HRTF よりも有意に小さい($p < 0.05$)ことが示された.

表 3.11 OIS と TCY の HRTF, best-matching HRTF の平均定位誤差

Subject	HRTF	Taget vertical angle (°)							Ave.
		0	30	60	90	120	150	180	
OIS	TCY	20.6	19.9	33.0	16.3	23.3	39.5	7.1	22.8
	best-matching	5.2	36.7	44.3	39.8	20.4	47.3	3.4	28.2
TCY	OIS	2.5	15.2	26.3	42.2	36.0	66.0	1.2	27.0
	best-matching	0.3	1.9	4.9	30.5	36.7	40.0	2.4	16.7
CKT	OIS	75.0	17.9	33.8	27.4	30.7	38.4	1.0	32.0
	TCY	172.1	10.2	18.5	12.3	46.2	71.2	1.3	47.4
	best-matching	2.1	29.2	30.8	21.3	31.9	17.3	3.5	19.5
MTZ	OIS	108.2	62.4	61.7	48.3	30.6	17.9	85.9	59.3
	TCY	178.6	25.7	40.2	32.0	22.8	29.4	70.6	57.0
	best-matching	3.3	24.8	27.9	38.5	20.6	13.1	0.5	18.4

表 3.12 OIS と TCY の HRTF, best-matching HRTF の前後誤判定率

Subject	HRTF	Target vertical angle (°)							Ave.
		0	30	60	90	120	150	180	
OIS	TCY	10	0	60		20	20	0	18.3
	best-matching	0	10	60		0	30	0	16.7
TCY	OIS	0	0	30		50	60	0	23.3
	best-matching	0	0	0		30	20	0	8.3
CKT	OIS	50	0	40		30	20	0	23.3
	TCY	100	0	20		100	90	0	51.7
	best-matching	0	0	30		20	0	0	8.3
MTZ	OIS	60	60	90		0	0	50	43.3
	TCY	100	10	80		0	0	50	40.0
	best-matching	0	10	50		20	0	0	13.3

表 3.13 非個人化 HRTF と best-matching HRTF の

前後誤判定率に対するカイ二乗検定の結果

Subject's best-matching HRTF	Other's HRTF	p-value
OIS	TCY	0.81
TCY	OIS	0.024
CKT	OIS	0.024
	TCY	2.2E-07
MTZ	OIS	2.7E-07
	TCY	9.6E-04

3.8 考察

従来の HRTF 個人化方法との比較を行う。本提案法では、受聴者の 6 箇所耳介形状を計測し、重回帰式を用いて N1, N2 周波数を推定し、それに最も近い best-matching HRTF をデータベースから選出した。このプロセスに必要な時間は約 3 分である。これは従来法より際立って速い。Middlebrooks *et al.* [59] は受聴者に適合した scale factor を見出すには 1 ブロックで 20 分要する聴覚実験が 1 から 3 ブロック必要であると述べている。Iwaya [54] はトーナメント方式の聴覚実験を提案しているが、32 の HRTF セットから受聴者に適したセットを選出するのに 15 分要すると報告している。この場合、データベースに含まれる HRTF が増えれば増えるほど聴覚実験が必要とする時間が増えることに留意しなければならない。

また、本提案法は作業者に特別な訓練も装置も要求しない。

3.9 結論

精度の高い音像制御を実現することを目的として、受聴者の耳介形状から HRTF の N1,N2 周波数を推定する方法を提案した。さらに、推定値に最も近い N1,N2 周波数をもつ best-matching HRTF を HRTF データベースから選出する方法を提案した。これらの提案法の有効性を検証するために音像定位実験を行った。それらの結果を以下にまとめる。

- (1) 日本人成人 54 耳の正面方向の N1,N2,P1 周波数を実測 HRTF から抽出した。個人差は、それぞれ 0.74, 0.71, 0.31 octave であった。
- (2) 54 耳の正面方向の N1,N2 周波数を目的変数、6 箇所 of 耳介形状を説明変数として重回帰分析を行った。N1,N2 の重相関係数はそれぞれ 0.81, 0.82 であった。
- (3) 4 人のナイーブな被験者の正面方向の N1,N2 周波数を 6 箇所 of 耳介形状から推定した。8 耳全てにおいて残差は JND 未満であった。
- (4) この 4 人の被験者を用いて上半球正中面で音像定位実験を行った。その結果、N1,N2 を推定した 0° (正面) では、best-matching HRTF の音像定位精度は被験者本人の HRTF と同等であった。
- (5) 180° (後方) においても、best-matchig HRTF の音像定位精度は被験者本人の HRTF と同等であった。
- (6) 他の上方の 5 方向では、幾人かの被験者の音像定位精度は本人の HRTF よりも低下した。
- (7) 提案した個人化プロセスの所要時間は約 3 分で、従来の方法よりも顕著に短い。また、耳介形状の計測は特別な訓練や設備を必要としない。

以上より、提案法により、0, 180° においては実測 HRTF と同等の精度で音像制御を提供できることを示した。しかし、上方においては一部の被験者において音像定位精度が低下する場合があります。検討の余地を残した。提案法は、推定した正面方向の N1,N2 周波数に基づいてデータベースから上半球正中面内の HRTF を選出するものである。したがって、正面方向の N1,N2 周波数だけでは、正中面内の HRTF の個人化精度に限界がある可能性がある。第 4 章では、上方における音像定位精度の向上を目指し、上方におけるスペクトラルキューの役割に関する検討を行い、今後推定すべきスペクトラルキューについて考察した。

さらに、提案法では、N1, N2 の周波数のみを推定したが、N1, N2 のレベルや尖鋭度についても推定できれば、各受聴者に適合する HRTF をパラメトリックに生成でき、

HRTF データベースが必要なくなる可能性がある。ただし、N1, N2 のレベルや尖鋭度の個人差の許容値については十分な検討は行われていない。また、著者らは、本章と同様の方法で、耳介形状から N1, N2 のレベルの推定を試みたが、高い精度で推定することはできなかった。耳介形状から N1, N2 のレベルや尖鋭度を推定するためには、今後、何らかの新しいアプローチが必要である。

第4章 頭部伝達関数の第2ピークが

正中面上方の音像定位精度に及ぼす影響

4.1 本章の目的

第3章における best-matching HRTF を用いた音像定位実験の結果、上方において一部の被験者では被験者本人の実測 HRTF に比べて best-matching HRTF の音像定位精度が低下する場合があった。本章では、上方の音像定位精度の向上のため、HRTF の第2ピークが上方の音像定位精度に及ぼす影響を検証した。

Hebrank and Wright [9]は上半球正中面 (-30°から+210°まで、30°間隔) に設置されたスピーカを用いて無響室で音像定位実験を行った。刺激は帯域幅が low-pass, high-pass, band-pass, band-stop filters で制限された白色雑音である。彼らは 16 kHz 以上もしくは 3.8 kHz 以下のスペクトルエネルギーは音像定位性能に影響しないこと、つまり正中面定位に利用されている特性は 3.8 から 16 kHz であることを示した。さらに彼らは、フィルタ処理なしと比較して、ある方向に 1.5 倍以上の回答が得られたフィルタをスペクトラルキューと定義し、前方のキューは低域側のカットオフ周波数が 4 kHz から 8 kHz の間にある 1 octave のノッチと 13 kHz 以上のエネルギーの上昇、上方のキューは 7 kHz から 9 kHz の間の 1/4 octave のピーク、そして後方のキューは 10 kHz から 12 kHz の間にある小さなピーク（このピークの高域側および低域側でのエネルギーの減衰を伴う）であると報告した。

Iida *et al.* [15]は上昇角音像定位をシミュレートするパラメトリック HRTF を提案した。パラメトリック HRTF は、音源の上昇角に依存しない 4 kHz のピーク[16]を下限周波数として、受聴者の実測 HRTF から抽出したノッチやピークから再構成される。ノッチとピークは周波数の順にラベルがつけられる (P1, N1, P2, N2, ... のように)。ノッチとピークは中心周波数、レベル、先鋭度でパラメトリックに表現される。彼らは上半球正中面において音像定位実験を行い以下のことを示した。

- (1) すべてのノッチとピークで再構成されたパラメトリック HRTF は被験者本人の実測 HRTF と同等の定位性能を有する。
- (2) 4 kHz 付近の第1ピーク(P1)と、P1周波数より高く最も周波数の低い2つのノッチ(N1, N2)だけで再構成されたパラメトリック HRTF は、前方と後方において

は被験者本人の実測 HRTF と同等の定位性能を有する。

- (3) しかし上方においては, N1, N2, P1 で再構成されたパラメトリック HRTF の音像定位性能は, 被験者によっては本人の実測 HRTF と比較して低下する。
- (4) N1 と N2 の周波数は音源の上昇角に強く依存するが, P1 周波数はほぼ一定で上昇角に依存しない。

これらの結果より, 彼らは N1 と N2 は, 少なくとも前方と後方では, スペクトラルキューとして重要な役割を果たしていると結論付けた。さらに彼らは, ヒトの聴覚システムが耳入力信号から N1 と N2 を分析するための参照情報として P1 を用いている可能性を指摘した。

Hebrank and Wright[9]は上方のスペクトラルキューは 7 kHz から 9 kHz の間の 1/4 octave のピークであると主張した。このピークは P2 と一致し, さらに Blauert[74]が提案した上方の方向決定帯域とも一致する。しかし, 狭帯域信号を用いて得られた方向決定帯域が広帯域信号に対するスペクトラルキュー の役割を果たすか否かは不明である。本章では, 以下の 2 つの目的を有する。

- 1) N1, N2, P1 に P2 を加え, N1, N2, P1, P2 で再構成されたパラメトリック HRTF が上方において実測 HRTF と同等の音像定位性能を有するか否かを検証する。
- 2) 正中面上方の音像定位における P2 の役割を解明する。

4.2 パラメトリック頭部伝達関数の再構成方法

4.2.1 頭部伝達関数の測定

無響室において, 22-24 歳の 3 人の男性被験者(MKI,OIS, OTK)および 1 人の女性被験者(YSD)の上半球正中面 7 方向 (30°間隔) の HRTF を測定した。

4.2.2 ノッチとピークの抽出

N1, N2, P1, P2 は耳介で生成されているため[19-21], 0°と 180°については被験者の頭部インパルス応答(HRIR)の初期部分から抽出した。そのアルゴリズムを以下に示す。

- (1) HRIR の絶対値が最大となるサンプルを検出する。
- (2) 検出したサンプルを中心として 4 次 96 サンプルの Blackman-Harris 窓で HRIR を切り出す。
- (3) 全ての要素をゼロとした 512 ポイントの配列を用意し, (2) で切り出した HRIR

を上書きする。ただし、最大サンプルを配列の 257 番目に合わせる

(4) FFT により 512 ポイントの配列の振幅スペクトルを求め、差分法で極大値および極小値をすべて検出する。

(5) 3 kHz 以上で最も周波数の低い 2 つの極大値をそれぞれ P1,P2 とし、P1 より高い帯域で最も周波数の低い 2 つの極小値を N1,N2 とする。

30°から 150°では、被験者によっては N1,N2 は浅くて不明瞭な場合があるので、以下のアルゴリズムにより N1,N2 を抽出した。

(1) FFT により 512 サンプルの HRIR の振幅スペクトルを求め、差分法で極小値をすべて検出する。

(2) Iida and Ishii [73]が報告した以下の回帰式を用いて N1,N2 周波数を算出する。彼らは、正面方向の N1,N2 周波数は被験者に強く依存するが、上昇角の関数としての N1,N2 周波数の振舞いは被験者間で共通であるとみなせると報告している。

$$f_{N1}(\beta) = 1.001 \times 10^{-5} \times \beta^4 - 6.431 \times 10^{-3} \times \beta^3 + 8.686 \times 10^{-1} \times \beta^2 - 3.265 \times 10^{-1} \times \beta + f_{N1}(0) \text{ [Hz]} \quad (4.1)$$

$$f_{N2}(\beta) = 1.310 \times 10^{-5} \times \beta^4 - 5.154 \times 10^{-3} \times \beta^3 + 5.020 \times 10^{-1} \times \beta^2 + 2.563 \times 10 \times \beta + f_{N2}(0) \text{ [Hz]} \quad (4.2)$$

ここで f_{N1} と f_{N2} はそれぞれ N1, N2 周波数[Hz]を表し、 β は上昇角[deg.]である。

(3) (1)で検出した極小値のうち、(2)で算出した N1, N2 周波数から 0.2 octave 以内で最も深い極小値をそれぞれ N1, N2 とする。その理由は、上昇角知覚における N1, N2 の丁度可知差(JND)は 0.1 から 0.2 octave と考えられる[51]からである。P1, P2 周波数は上昇角に依存しないので、30°から 150°の P1, P2 は 0°と同様とした。

4.2.3 パラメトリック頭部伝達関数の生成

パラメトリック HRTF は、2 次の IIR フィルタで作成したノッチとピークを重ね合わせることにより生成した。実測 HRTF, N1,N2,P1 で再構成したパラメトリック HRTF (以降, N1N2P1 と呼ぶ), N1, N2, P1, P2 で再構成したパラメトリック HRTF (以降, N1N2P1+P2 と呼ぶ) の例を図 4.1 に示す。パラメトリック HRTF は再構成したノッチとピークを精度よく再現し、その他の帯域のスペクトルは平坦である。

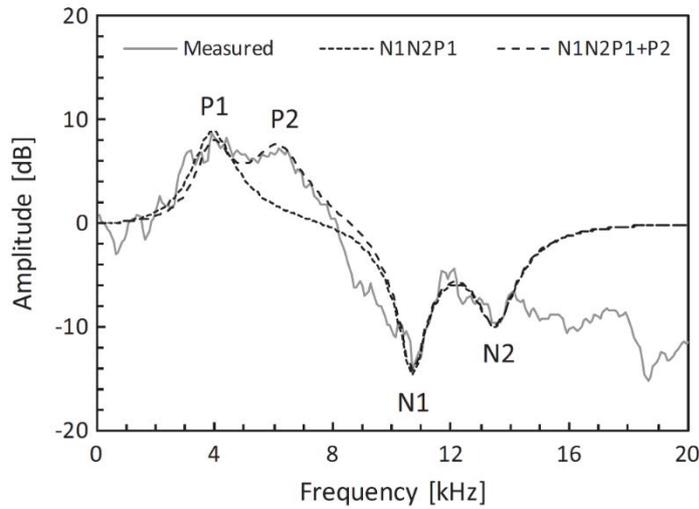


図 4.1 上昇角 90°のパラメトリック HRTF(N1N2P1, N1N2P1+P2)の例.

4.3 音像定位実験

4.3.1 音像定位実験(1)

N1N2P1+P2 が上方において実測 HRTF と同等の音像定位精度を有するか否かを検証するため、上半球正中面において音像定位実験(1)を行った。

A 実験方法

被験者は 4.2.1 項で HRTF の測定を行った 4 人(MKI, OIS, OTK, YSD)である。HRTF は以下の 3 種類を用いた。

- (1) 被験者本人の実測 HRTF
- (2) 被験者本人の N1N2P1
- (3) 被験者本人の N1N2P1+P2

実測 HRTF と 2 種類のパラメトリック HRTF の音像定位実験は別々に行った。Møller *et al.*[26]は 2 つの条件,

- (1) 1 人の被験者の HRTF をランダム化した条件
- (2) 数人の HRTF をランダム化した条件

を比較し、同じ刺激群を分離して実験しても混合しても有意な差はないことを示している。その他の実験方法は第 3 章における音像定位実験と同様である。

B 実験結果

a) 回答の分布

実験結果を図 4.2 に示す。実測 HRTF では、ほとんどの回答は目標方向付近に分布した。ただし、被験者 OTK の回答は、60, 120, 150°の目標方向では 90°付近に分布した。被験者 YSD の 0°でやや上方に、90°でやや後方にシフトした。

N1N2P1 では、被験者 MKI および OIS の回答の分布は、実測 HRTF と同様であった。ただし、被験者 MKI の 120°、被験者 OIS の 90°で定位精度が低下した。被験者 OTK の回答は、60, 90, 120°では、0°および 90°周辺に分布した。

N1N2P1+P2 では、すべての被験者において、一定の目標方向において定位精度が向上した。被験者 MKI については、120, 150°において定位精度が改善された。被験者 OIS については、90°において定位精度が改善された。ただし、150°では上方にシフトしたが、この理由は不明である。被験者 OTK については、N1N2P1 において観測された分離した回答の分布は、60, 90, 120°では観測されなかった。N1N2P1+P2 に対する回答は、実測 HRTF の回答と同等であった。被験者 YSD については、0°において定位精度が改善された。

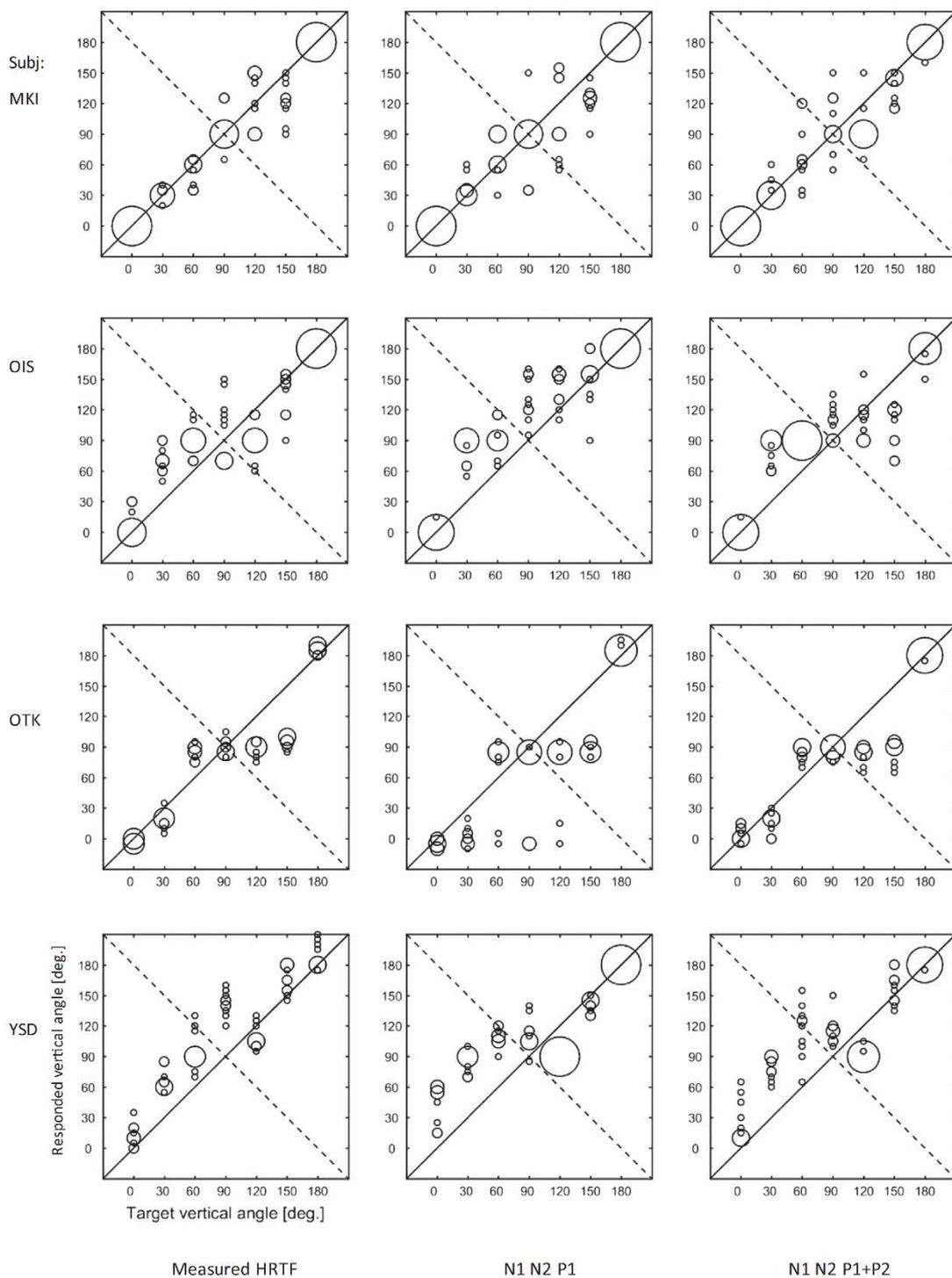


図 4.2 実験結果：回答の分布

b) 平均上昇角定位誤差

各 HRTF と各目標方向の平均上昇角定位誤差を求めた (表 4.1). ここで, 平均上昇角定位誤差は目標の上昇角と回答した上昇角の差の絶対値の平均と定義する. 実測 HRTF の平均上昇角誤差は, Carlile *et al.* [76]や Majdak *et al.* [46]の報告と同様, 水平面上の $0^\circ, 180^\circ$ では小さく, 水平面から上昇するほど大きくなる傾向がある.

表 4.1 平均上昇角定位誤差

Subject	HRTF	Target vertical angle ($^\circ$)							Ave.
		0	30	60	90	120	150	180	
MKI	Measured	0.5	3.5	10.0	10.0	23.0	27.1	0.4	10.6
	N1N2P1	0.8	6.9	15.7	17.8	39.0	28.3	1.0	15.6
	N1N2P1+P2	0.3	5.2	22.2	21.1	30.1	15.4	2.8	13.9
OIS	Measured	8.7	40.4	30.5	28.1	29.7	16.4	0.5	22.0
	N1N2P1	2.3	51.1	31.6	41.8	23.3	18.5	0.7	24.2
	N1N2P1+P2	2.2	50.1	30.1	19.3	16.6	46.6	3.7	24.1
OTK	Measured	2.8	13.0	24.6	6.2	32.3	55.5	5.6	20.0
	N1N2P1	5.1	28.8	31.6	31.5	50.5	61.9	6.4	30.8
	N1N2P1+P2	6.3	13.7	23.5	5.1	37.0	64.2	1.4	21.6
YSD	Measured	12.4	36.7	35.8	52.0	14.3	16.1	13.0	25.8
	N1N2P1	44.4	54.1	49.0	22.9	30.0	9.7	0.8	30.2
	N1N2P1+P2	27.4	48.4	55.1	25.2	28.3	13.8	0.8	28.4
Average across subjects	Measured	6.1	23.4	25.2	24.1	24.8	28.8	4.9	19.6
	N1N2P1	13.2	35.2	32.0	28.5	35.7	29.6	2.2	25.2
	N1N2P1+P2	9.0	29.3	32.7	17.7	28.0	35.0	2.2	22.0

表 4.2 Tukey の多重比較検定の結果

Comparison between	Target vertical angle ($^\circ$)						
	0	30	60	90	120	150	180
Measured and N1N2P1		*			**		
Measured and N1N2P1+P2							
N1N2P1 and N1N2P1+P2							

** $p < 0.01$

* $p < 0.05$

N1N2P1 の平均上昇角誤差は, ほとんどの目標方向および被験者において実測 HRTF よりも大きかった. 特に, $30, 60, 90, 120^\circ$ では, 実測 HRTF よりも大きい傾向にある.

N1N2P1+P2 の平均上昇角誤差は, 多くの目標方向および被験者において, N1N2P1 よりも小さい. 特に, N1N2P1 に P2 を加えることで, 被験者 YSD の 0° , OIS と OTK の 90° , MKI, OIS, OTK の 120° , MKI の 150° で誤差が減少した. しかし, N1N2P1+P2 の

誤差は、OIS の 150°で N1N2P1 の誤差よりも大きかった。これは理由が不明であるが OIS の回答が上方にシフトしたためである。

全被験者の平均上昇角誤差は、0, 30, 90, 120°では、N1N2P1 に P2 を加えることで減少した。特に、平均上昇角誤差は、90°で 10.8°減少した。さらに、7つの目標方向において、N1N2P1+P2 と実測 HRTF の平均上昇角誤差の差は 10°未満であった。

次に、実測 HRTF, N1N2P1, N1N2P1+P2 の平均定位誤差の間で統計的有意差があるか否かについて、Tukey の多重比較検定を行った。表 4.2 に結果を示す。N1N2P1 の平均上昇角誤差は、30°($p<0.05$)と 120°($p<0.01$)で実測 HRTF よりも有意に大きかった。一方、実測 HRTF と N1N2P1+P2 の間、または N1N2P1 と N1N2P1+P2 の間では、いずれの目標方向においても有意差はなかった。

これらの結果は、N1N2P1+P2 は、上半球正中面における 7つの目標方向のいずれにおいても、実測 HRTF と同等の定位精度を有するが、N1N2P1 では 30, 120°で実測 HRTF よりも定位精度が低下することを示している。

c) 前後誤判定率

4名の被験者の平均前後誤判定率を表 4.3 に示す。実測 HRTF, N1N2P1, N1N2P1+P2 の前後誤判定率は、0, 180°では 0%であった。しかし、他の 4つの目標方向では、N1N2P1 と N1N2P1+P2 の前後誤判定率は、実測 HRTF よりも高かった。N1N2P1 の前後誤判定率は、60°を除いて実測 HRTF の約 2 倍であった。N1N2P1+P2 の前後誤判定率は、N1N2P1 よりもわずかに低い傾向がある。

実測 HRTF, N1N2P1, N1N2P1+P2 の前後誤判定率の間で統計的有意差があるか否かを検証するため、カイ二乗検定を行った。その結果、実測 HRTF と N1N2P1 の間、および実測 HRTF と N1N2P1+P2 の間の 120°で統計的有意差が認められた。いずれの目標方向においても、N1N2P1 と N1N2P1+P2 の間には有意差は認められなかった。

表 4.3 前後誤判定率 [%]

HRTF	Target vertical angle (°)						
	0	30	60	90	120	150	180
Measured	0	5	35	-	23	10	0
N1N2P1	0	13	45	-	50	20	0
N1N2P1 + P2	0	5	43	-	50	15	0

d) 頭内定位率

いずれの HRTF においても頭内定位は発生しなかった。

4.3.2 音像定位実験(2)

音像定位実験 1 では, N1N2P1 に P2 を加えることにより正中面上方の音像定位精度が向上することを示した. 音像定位実験(2)の目的は, P2 が正中面上方の音像定位で担う役割を解明することである.

A 実験方法

実験では, 被験者本人の 90°の

(1) P1

(2) P2

(3) P1P2

で再構成したパラメトリック HRTF を用いた. その他の条件は実験 1 と同様である. パラメトリック HRTF の一例を図 4.3 に示す.

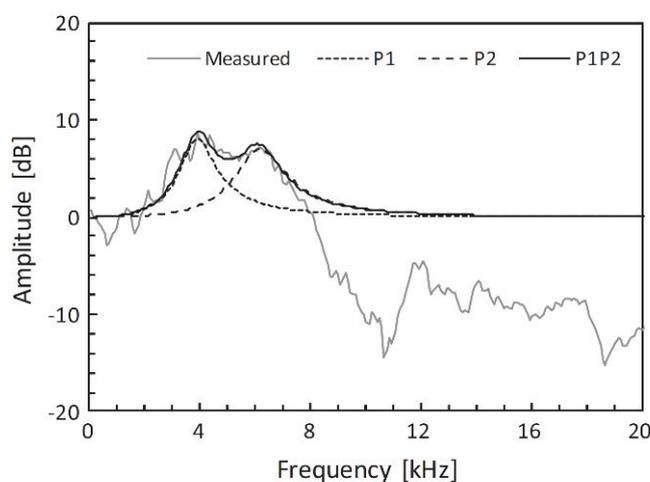


図 4.3 上昇角 90°における実測 HRTF とパラメトリック HRTF(P1, P2, P1P2)の一例

B 実験結果

a) 回答の分布

P1, P2, P1P2 に対する回答分布を図 4.4 に示す. 音像定位実験 1 で求めた目標方向 90°における実測 HRTF, N1N2P1, N1N2P1+P2 の回答分布も併せて示す. N1N2P1+P2 の回答分布は, MKI でやや分散が大きくなるものの, 実測 HRTF のそれと同様である.

一方, P1, P2, P1P2 に対する回答は被験者内でほぼ同様に分布した. MKI と OIS は前方から後方, OTK は上方から後方, YSD は斜め後ろから後方に広く分布した. いずれの被験者においても, 上方に回答が集中することはなかった.

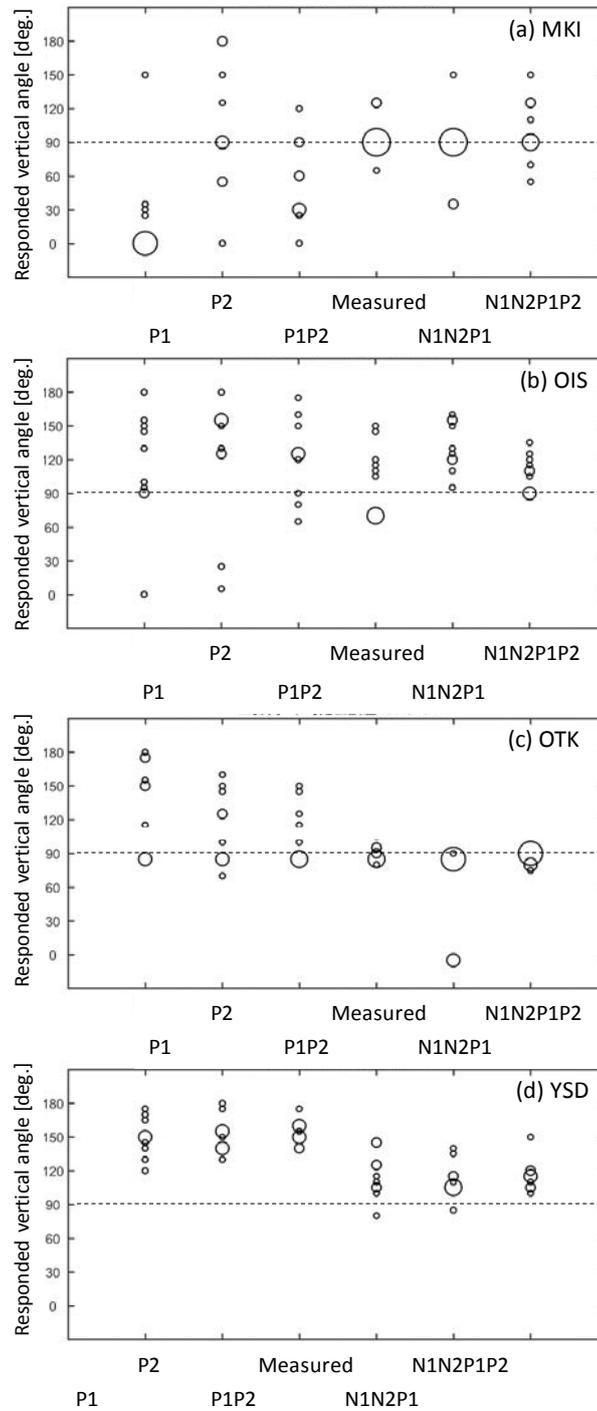


図 4.4 回答の分布

b) 平均定位誤差

全被験者で平均した上昇角定位誤差を図 4.5 に示す。N1N2P1+P2 は実測 HRTF と同等であった。しかし、P1、P2、P1P2 は実測 HRTF と比較して大きい値であった。

実測 HRTF とパラメトリック HRTF の間で有意差があるか否かを検証するため、Tukey の多重比較検定を行った。表 4.4 に結果を示す。P1、P2、P1P2 の平均上昇角誤差は実測 HRTF よりも有意に大きいが、N1N2P1+P2 の平均上昇角誤差は実測 HRTF と有意差は認められなかった。これらの結果は、スペクトラルピークはそれ自体が上方の定位の手掛かりとして十分ではないことを示している。

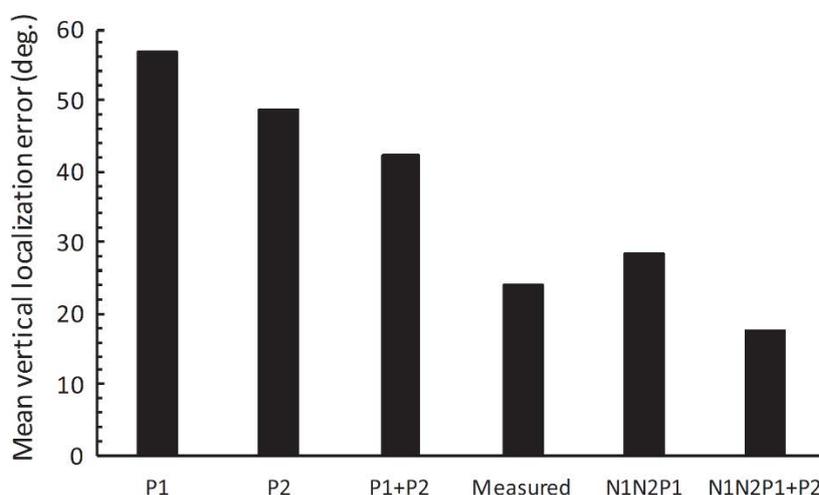


図 4.5 平均上昇角誤差

表 4.4 平均上昇角誤差に対する Tukey の多重比較検定の結果

	P1	P2	P1P2	Measured	N1N2P1	N1N2P1 + P2
P1	—					
P2		—				
P1P2			—			
Measured	**	**	*	—		
N1N2P1	**	**			—	
N1N2P1 + P2	**	**	**			—

** p < 0.01

* p < 0.05

4.4 考察

図 4.6 に目標方向 90°において、N1N2P1 で上方と前方に回答し、N1N2P1+P2 では上方に回答した被験者 OTK の N1,N2,P1,P2 の周波数とレベルの関係を示したものである。白丸は 0°, 黒丸は 90°の N1,N2,P1,P2 を示す。2 本の破線は、Moore *et al.*[18]が中心周波数 8 kHz, 帯域幅が中心周波数の 25%のノッチを用いて求めた 3 人の被験者のノッチ検知閾の最大値と最小値を示している。すなわち、ノッチレベルが -9 dB 以上ではすべての被験者がノッチを検知できず、-20 dB 以下ではすべての被験者が検知した。0°の N1,N2 のレベルはともに検知閾を超えている。しかし、90°の N1 は検知できないレベルであり、N2 は被験者により検知の可否が分かれるレベルである。周波数軸上でみると、90°ではノッチとピークは周波数の低い方から P1, P2, N1, N2 の順で並んでいる。P2 を再現しない場合、P1 は N1 から離れた周波数にあるため対比効果は期待できない。しかし P2 を再現すると、P2 からみた N1 の相対的なレベルは -14.7 dB となる。これは被験者によってはノッチの検知が可能なレベルである。以上より、P2 それ自体はスペクトラルキューではないが、N1 を強調することにより、正中面上方の音像定位精度を向上させる重要な役割を果たしていると考えられる。

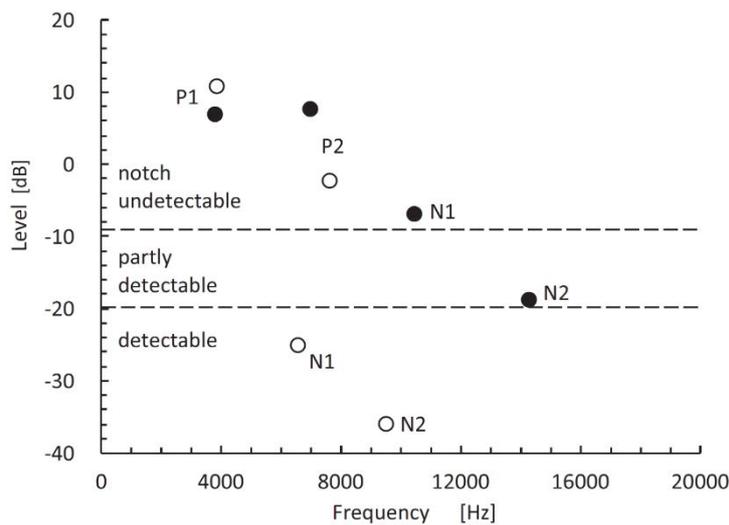


図 4.6 N1, N2, P1, P2 の周波数とレベルの関係

4.5 結論

本章では、N1N2P1 に P2 を加えることで上方の定位精度が向上するか否かを検証するため、4名の被験者を用いて上半球面内7方向を目標方向とした音像定位実験を行った。

その結果、以下のことを示した。

- 1) N1N2P1 の平均上昇角誤差は、実測 HRTF のそれよりも 30, 120°で有意に大きかった。
- 2) N1N2P1 に P2 を加えることにより、正中面の 0, 30, 90, 120°で平均上昇角誤差が減少した。また、N1N2P1+P2 と実測 HRTF の平均上昇角誤差は、7つの目標方向において 10°未満であり、統計的有意差は認められなかった。
- 3) 実測 HRTF, N1N2P1, N1N2P1+P2 の前後誤判定率の間では、実測 HRTF と N1N2P1, 実測 HRTF と N1N2P1+P2 の 120°を除き、いずれの目標方向においても有意差は認められなかった。
- 4) P1,P2,P1P2 を再現しても音像は上方に知覚されなかった。言い換えれば、これらのスペクトラルピークは上方のスペクトラルキューではない。
- 5) 0°の N1 と N2 のレベルはともにノッチの検知閾を超えるが、90°の N1 は検知できない。しかし、P2 を加えることにより、P2 からみた N1 の相対レベルは検知可能となる。
- 6) 以上より、P2 は N1 を強調することにより、正中面上方の音像定位精度を向上する重要な役割を果たしていると考えられる。

以上より、上半球正中面内全体において、実測 HRTF と同等の定位精度を提供できる最小構成は、N1, N2, P1, P2 であることが示唆された。

本章の結果から、第3章と同様の方法で耳介形状から P2 も推定し、推定した N1, N2, P2 に基づいて適合する HRTF を選出することで、上方における音像定位精度の向上が期待できる。今後、受聴者の耳介形状から推定した N1, N2, P2 周波数に基づいて適合する HRTF(best-matching HRTF)を選出する方法について検討する余地が残されている。

第5章 総括

5.1 結論

本論文では、3次元音像制御の実用化のため、耳入力信号再現方式における HRTF の個人差に関する問題を解決するため、受聴者の頭部形状から HRTF の個人化を行う方法を提案し、その有効性を検証した。本研究では、まず、頭部形状から両耳間差キューを推定した。次に、耳介形状からスペクトラルキューを推定し、それに最も近い HRTF をデータベースから選出する方法を提案した。また、上方における音像定位精度の向上を目的とし、HRTF の第2ピークが上方の音像定位精度に及ぼす影響を検証した。

各章で得られた結果をまとめる。第2章では、受聴者の頭部形状から左右方向知覚の手掛かりである両耳間時間差(ITD)および両耳間レベル差(ILD)を個人化する方法を提案し、その有効性を検証した。

- 1) 各方向の ITD を目的変数、10箇所頭部形状を説明変数とした重回帰分析を行った。その結果、重相関係数は 0.34–0.79 であり、全ての方向において 5% で有意な相関関係が認められた。
- 2) 4名のナイーブな被験者を用いて提案法の推定精度を検証した。その結果、各方向の平均方位角誤差は 1.1–3.3° であり方向知覚の弁別閾と同等であった。ただし、各被験者の各方向の方位角誤差をみると、一部において方向知覚の弁別閾を超える場合があり、最大推定誤差は 7.1° (被験者 B の 60, 210°) であった。また、提案法の平均方位角誤差は 2.3° であり、従来法を用いた場合の平均方位角誤差よりも 7.3° 小さかった。
- 3) 10箇所頭部形状を説明変数、1/3 octave band (250–16000 Hz) の ILD を目的変数とした重回帰分析を行った結果、重相関係数の平均値は 0.58 であり、残差の絶対値の平均値は 1.67 dB であった。
- 4) 4名のナイーブな被験者を用いて重回帰式の推定精度を検証した。その結果、800 Hz 以下では ILD を弁別閾内 (約 1 dB) の誤差で推定可能であった。ただし、1000 Hz 以上の帯域では推定誤差は 2–9 dB であった。

以上より、頭部の左右非対称性を考慮し、方向ごとに重回帰モデルを構築した結果、ITD および 800 Hz 以下の ILD は弁別閾内の精度で推定できることを示した。しかし、1000 Hz 以上の帯域では弁別閾を超え、提案法の推定精度が従来法のそれに比べて改善

されなかった。この要因として、提案法では、従来法で用いられている鼻と耳介に関するパラメータを除外していたことが挙げられ、検討の余地を残した。

第3章では、受聴者の頭部形状から前後・上下方向知覚の手掛かりであるスペクトラルキューを耳介形状から推定し、それに最も近い HRTF をデータベースから選出することで正中面内の HRTF を個人化する方法を提案し、その有効性を検証した。

- 1) 54 耳の正面方向の N1, N2 を抽出した。N1, N2 周波数の個人差は、それぞれ 0.74, 0.71 octave であった。
- 2) 54 耳の正面方向の N1, N2 周波数を目的変数、6 箇所の耳介形状を説明変数とした重回帰分析を行った。N1, N2 周波数の重相関係数はそれぞれ 0.81, 0.82 であった。
- 3) 4 名のナイーブな被験者を用いて正面方向の N1, N2 周波数を推定した。8 耳全てにおいて残差は弁別閾以内であった。
- 4) この 4 名の被験者を用いて上半球正中面 7 方向を目標方向とした音像定位実験を行った。その結果、N1, N2 周波数を推定した 0° (正面) では、best-matching HRTF の音像定位精度は被験者本人の HRTF と同等であった。
- 5) 180° (後方) においても、best-matching HRTF の音像定位精度は被験者本人の HRTF と同等であった。
- 6) 他の上方の 5 方向では、幾人かの被験者の音像定位精度は本人 HRTF よりも低下した。

以上より、提案法により、上昇角 0, 180° においては実測 HRTF と同等の精度で音像制御を提供できることを示した。しかし、上方においては一部の被験者において音像定位精度が低下する場合があります。検討の余地を残した。

第4章では、上方の音像定位精度の向上を目的とし、HRTF における第2ピーク (P2) が上方の音像定位精度に及ぼす影響を音像定位実験により検証した。

- 1) N1N2P1 の平均上昇角誤差は、実測 HRTF のそれよりも 30, 120°で有意に大きかった。
- 2) N1N2P1 に P2 を加えることにより、正中面の 0, 30, 90, 120°で平均上昇角誤差が減少した。また、N1N2P1+P2 と実測 HRTF の平均上昇角誤差は、7つの目標方向において 10°未満であり、統計的有意差は認められなかった。
- 3) 実測 HRTF, N1N2P1, N1N2P1+P2 の前後誤判定率の間では、実測 HRTF と N1N2P1, 実測 HRTF と N1N2P1+P2 の 120°を除き、いずれの目標方向においても有意差は認

められなかった。

- 4) ピークのみ(P1, P2, P1+P2)を再現しても音像は上方に知覚されなかった。言い換えれば、これらのスペクトラルピークは上方のスペクトラルキューではない。
- 5) 0°の N1 と N2 のレベルはともにノッチの検知閾を超えるが、90°の N1 は検知できない。しかし、P2 を加えることにより、P2 からみた N1 の相対レベルは検知可能となる。
- 6) 以上より、P2 は N1 を強調することにより、正中面上方の音像定位精度を向上する重要な役割を果たしていると考えられる。

以上より、上半球正中面内全体において、実測 HRTF と同等の定位精度を提供できる最小構成は、N1, N2, P1, P2 であることが示唆された。今後、受聴者の耳介形状から推定した N1, N2, P2 周波数に基づいて適合する HRTF を選出する方法について再検討する余地が残されている。

5.2 今後の展望

- 1) 提案法では正面方向の N1, N2 周波数を推定し、その周波数に最も近い周波数を持つ HRTF を best-matching HRTF として選出し、他の方向の HRTF については正面方向の best-matching HRTF の提供者のものを用いた。今後は、耳介形状から正面方向以外の N1, N2 周波数を推定する方法を確立する必要がある。
- 2) 正面方向を含めた N1, N2 のレベル・尖鋭度の推定が可能となれば、best-matching HRTF の選出の際に、推定した N1, N2 のレベル・尖鋭度を考慮することで音像定位精度の向上が期待できる。また、N1, N2 の周波数、レベル、尖鋭度が推定できれば、HRTF データベースから選出することなく、信号処理により HRTF を生成できる可能性がある。しかし、著者らが提案法と同様の方法で N1, N2 のレベルの推定を試みたが、高い精度で推定することはできなかった。N1, N2 のレベルや尖鋭度を推定するためには、今後、何らかの新しいアプローチが必要であると考えられる。
- 3) 3次元音像制御の実用化に向け、HRTF の推定に用いる受聴者の頭部形状や耳介形状をカメラ等で簡易的に計測する方法を確立する必要性がある。ただし、HRTF の個人化において、耳介形状の奥行等の情報が必要であるため、立体的な画像処理が必要である。

このような研究課題が解決されれば、Blauert が示した以下のシナリオの通り、誰もが

個人化した HRTF を容易に得られる時代となると考えられる[77].

“A person who enters a multimedia shop is scanned by a camera and some instants later his/her individual HRTF set is ready to be sold for the use in advanced 3D applications”

参考文献

- [1] 飯田 一博, 森本 政之, “空間音響学,” コロナ社, pp.1-5, 東京, 2010
- [2] M. Morimoto and Y. Ando, “On the simulation of sound localization,” Journal of the Acoustical Society of Japan (E), 1(3), pp.167-174, 1980
- [3] H. Møller, “Fundamentals of Binaural Technology,” Applied Acoustics, 36(3/4), pp.171-218, 1992
- [4] L. Rayleigh, “Acoustical observations,” Philosophical Magazine, 3, 6th series, pp.456-464, 1877
- [5] L. Rayleigh, “On our perception of sound direction,” Philosophical Magazine. 13, 6th series, pp.214-232, 1907
- [6] F. E. Toole and Bruce McA Sayers: Lateralization judgements and the nature of binaural acoustic images, The Journal of the Acoustical Society of America, 37, 319-324, 1965
- [7] 飯田 一博, 森本 政之, “空間音響学,” コロナ社, pp.6-21, 東京, 2010
- [8] B. M. Sayers, “Acoustic-image lateralization judgement with binaural tones,” The Journal of the Acoustical Society of America, 36, pp.923-926, 1964
- [9] J. Hebrank and D. Wright, “Spectral cues used in the localization of sound sources on the median plane,” The Journal of the Acoustical Society of America, 56, pp.1829-1831, 1974
- [10] 森本 政之, 斉藤 明博, “音の正中面定位について—刺激音の周波数範囲と強さの影響について—,” 音響学会聴覚研資料, H-40-1, 1977
- [11] R. A. Butler and K. Belendiuk, “Spectral cues utilized in the localization of sound in the median sagittal plane,” The Journal of the Acoustical Society of America, 61, pp.1264-1269, 1977
- [12] S. Mehrgardt and V. Mellert, “Transformation characteristics of the external human ear,” The Journal of the Acoustical Society of America., 61, pp.1567-1576, 1977
- [13] A. Kulkarni and H. Steven Colburn, “Role of spectral detail in sound source localization,” Nature, 396, pp.747-749, 1998
- [14] F. Asano, Y. Suzuki and T. Stone, “Role of spectral cues in median plane

- localization,” *The Journal of the Acoustical Society of America.*, 80, pp.159-168, 1990
- [15] K. Iida, M. Itoh, A. Itagaki, and M. Morimoto, “Median plane localization using a parametric model of the head-related transfer function based on spectral cues,” *Applied Acoustics*, 68, pp.835-850, 2007
- [16] E. A. G. Shaw and R. Teranishi, “Sound pressure generated in an external-ear replica and real human ears by a nearby point source,” *The Journal of the Acoustical Society of America.* 44, 240–249, 1968
- [17] P. X. Zhang and W. M. Hartmann, “On the ability of human listeners to distinguish between front and back”, *Hearing Research*, 260, pp.30-46, 2010
- [18] B. C. J. Moore, S. R. Oldfield, and G. J. Dooley, “Detection and discrimination of spectral peaks and notches at 1 and 8 kHz,” *The Journal of the Acoustical Society of America.*, 85, pp.820-836, 1989
- [19] E. A. G. Shaw, “Acoustical features of the human ear, in *Binaural and Spatial Hearing in Real and Virtual Environments*,” R. H. Gilkey and T. R. Anderson, Edits. (Erbaum, Mahwah, New Jersey), pp 25-47, 1997
- [20] H. Takemoto, P. Mokhtari, H. Kato, Ryouichi Nishimura and Kazuhiro Iida, “Mechanism for generating peaks and notches of head-related transfer functions in the median plane,” *The Journal of the Acoustical Society of America.*, 132, pp.3832-3841, 2012
- [21] Y. Kahana and P. A. Nelson, “Numerical modeling of the spatial acoustic response of the human pinna,” *Journal of Sound and Vibration*, 292, pp.148-178, 2006
- [22] P. Mokhtari, H. Takemoto, R. Nishimura and H. Kato “Comparison of simulated and measured HRTFs: FDTD simulation using MRI head data,” *123rd Audio Engineering society*, 7240, pp.1-12, 2007
- [23] P. Mokhtari, R. Nishimura and H. Takemoto, “Toward HRTF personalization: An auditory perceptual evaluation of simulated and measured HRTFs,” *14rd International Conference on Auditory Display*, 2008
- [24] V. C. Raykar, R. Duraiswami, and B. Yegnanarayana, “Extracting the

- frequencies of the pinna spectral notches in measured head-related impulse responses,” *The Journal of the Acoustical Society of America.*, 118, pp.364-374
2005
- [25] 蒲生 直和, 石井 要次, 飯田 一博 “仰角知覚のスペクトラルキューの形成における初期頭部インパルス応答の寄与,” *日本音響学会講演論文集(春)*, pp.517-520, 2011
- [26] H. Møller, D. Hammershøi, C. B. Jensen and M. F. Sørensen, “Transfer characteristics of headphones measured on human ears,” *Journal of the Audio Engineering Society*, 43, pp.203–217, 1995
- [27] F. L. Wightman and D. J. Kistler, “Headphone simulation of free-field listening. II: Psychophysical validation,” *The Journal of the Acoustical Society of America.*, 85, pp.868-878, 1989
- [28] M. R. Schroeder and B. S. Atal, “Computer simulation of sound transmission in rooms,” *IEEE international convention record, Part 7*, pp.150–155, 1963
- [29] P. Damaske and V. Mellert, “Ein Verfahren zur richtungstreuen Schallabbildung des oberen Halbraumes über zwei Lautsprecher,” *Acustica*, 22, pp.153– 162, 1969/70
- [30] P. Damaske, “Head-related two-channel stereophony with loudspeaker reproduction,” *The Journal of the Acoustical Society of America.*, 50, pp.1109–1115 1971
- [31] 浜田 晴夫, “基準的收音・再生を目的とする Orthostereophonic System の構成,” *音響学会誌*, 39, pp.337– 348,1983
- [32] O. Kirkeby, P. A. Philip A. Nelson and H. Hamada, “Local sound field reproduction using two closely spaced loudspeakers,” *The Journal of the Acoustical Society of America.*, 104, pp.1973–1981, 1998
- [33] T. Takeuchi, P. A. Nelson, O. Kirkeby and H. Hamada, “Robustness of the “Stereo Dipole” to misalignment of head position,” *AES 102nd Convention, Preprint 4464, Munich, Germany*, 1997
- [34] 飯田 一博, 森本 政之, *空間音響学*, コロナ社, pp. 118–122, 東京, 2010
- [35] T. Takeuchi and P. A. Nelson, “Optimal source distribution for binaural synthesis over loudspeakers,” *The Journal of the Acoustical Society of America.*, 112, pp.2786–

2797, 2002

- [36]伊藤 毅 “音響工学原論,” コロナ社, pp. 587 -589, 東京, 1973
- [37]及川 靖広, 馬場 貞如, 山崎 芳男, “人間の聴感を考慮した波面合成法による音場制御と音場再現,” 日本バーチャルリアリティ学会論文集, 5(3), pp.919-926, 2000
- [38]イエンス ブラウエルト, 森本 政之, 後藤敏幸, “空間音響,” pp. 193-195, 鹿島出版会, 東京, 1986
- [39]伊勢 史郎, “キルヒホッフ-ヘルムホルツ積分方程式と逆システム理論に基づく音場制御の原理,” 日本音響学会誌, 53(9), pp.706-713, 1997
- [40]D. H. Cooper and T. Shiga, “Discrete-Matrix Multichannel Stereo,” *Journal of Audio Engineering Society*, 20, pp.346-360, 1972
- [41]M. A. Gerzon, “Periphony, “Periphony: With-height sound reproduction,” *Journal of Audio Engineering Society*, 21, pp.2 -10, 1973
- [42]岩谷 幸雄, 岡本 拓磨, トレビーニョホルヘ, 鈴木陽一, “球面調和解析による音場表現 : 高次アンビソニックス技術の可能性, ” 日本音響学会誌, 67(11), pp.544-549, 2011
- [43]森川 大輔, 石井 要次, 高道 慎之介, Jorge Trevino, “立体音再生技術の歴史,” 第4回音響セミナー, 2013.
- [44]E. M. Wenzel, M. Arruda, D. J. Kistler and F. L. Wightman, “Localization using nonindividualized head-related transfer functions,” *The Journal of the Acoustical Society of America.*, 94, pp.111-122, 1993
- [45]X. Yan, K. Iida, and Y. Ishii, "Comparison in frequencies of spectral peaks and notches and anthropometric of pinnae between HRTF databases," 電子情報通信学会, pp.43-48, 2014
- [46]P. Majdak, M. J. Goupell, and B. Laback, “3-D localization of virtual sound sources: Effects of visual environment, pointing method and training,” *Atten Percept Psychophysics*, 72, pp.454-469, 2010
- [47]V. R. Algazi, R. O. Duda, D. M. Thompson and C. Avendano, “The CIPIC HRTF database,” *IEEE. Workshop on Application of Signal Processing to Audio and Acoustics*, pp.99-102, 2001
- [48]Spatial Hearing Laboratory – 千葉工業大学飯田研究室<<http://www.iida-lab.it>>

chiba.ac.jp/>, (最終アクセス日 : 2017/11/29)

- [49] listen hrtf database – Ircam, <<http://recherche.ircam.fr/equipes/salles/listen>>, (最終アクセス日 : 2017/11/29)
- [50] K. Watanabe, Y. Iwaya, Y. Suzuki, S. Takane and S. Sato, “Dataset of head-related transfer functions measured with a circular loudspeaker array,” *Acoust.Sci. & Tech.*, 35, pp.159-165, 2014
- [51] K. Iida and Y. Ishii, “Individualization of the head-related transfer functions in the basis of the spectral cues for sound localization,” in *Principles and Applications of Spatial Hearing*, edited by Y. Suzuki, D. Brungard, Y. Iwaya, K. Iida, D. Cabrera, and H. Kato (World Scientific, Singapore), pp. 159–178, 2011
- [52] M. D. Burkhard and R. M. Sachs, “Anthropometric manikin for acoustic research,” *The Journal of the Acoustical Society of America.*, 58, pp.214-222, 1975
- [53] B. U. Seeber and H. Fastl, “Subjective selection of non-individual head-related transfer functions,” *Proceeding of the 2003 International Conference on Auditory Display*, Boston, MA, USA, July 6-9, pp.259-262, 2003
- [54] Y. Iwaya, “Individualization of head-related transfer functions with tournament-style listening test: Listening with other’s ears,” *Acoust. Sci. & Tech.* 27, pp.340–343, 2006
- [55] D. J. Kistler and F. L. Wightman, “A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction,” *The Journal of the Acoustical Society of America.* 91, pp.1637-1647, 1992
- [56] J. C. Middlebrooks and D. M. Green, “Observations on a principal components analysis of head-related transfer functions,” *The Journal of the Acoustical Society of America.* 92, pp.597–599, 1992
- [57] J. C. Middlebrooks, “Individual differences in external-ear transfer functions reduced by scaling in frequency,” *The Journal of the Acoustical Society of America.* 106, pp.1480–1492, 1999
- [58] J. C. Middlebrooks, “Virtual localization improved by scaling nonindividualized external-ear transfer functions in frequency,” *The Journal of the Acoustical*

- Society of America. 106, pp.1493–1510, 1999
- [59] J. C. Middlebrooks, E. A. Macpherson, and Z. A. Onsan, “Psychophysical customization of directional transfer functions for virtual sound localization,” *The Journal of the Acoustical Society of America*. 108, 3088–3091, 2000
- [60] D. Y. N. Zotkin, J. Hwang, R. Duraiswami, and L. S. Davis, “HRTF personalization using anthropometric measurements,” in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, pp. 157–160, 2003
- [61] V. R. Algazi, C. Avendano, R. O. Duda, “Estimation of a Spherical-Head Model from Anthropometry,” *J. Audio Eng. Soc.* 49(6), pp.472-479, 2001
- [62] 渡邊 貫治, 岩谷 幸雄, 行場 次朗, 鈴木 陽一, 高根 昭一, “身体特徴量に基づく両耳間時間差の予測に関する検討,” *日本バーチャルリアリティ学会論文誌*, 10(4), pp.609-617, 2005
- [63] 黒澤 明, 都木 徹, 山口 善司, “頭部伝達関数と方向弁別能力について,” *日本音響学会誌*, 38(3), pp.145-151, 1982
- [64] K. Watanabe, K. Ozawa, Y. Iwaya, Y. Suzuki, and K. Aso, “Estimation of interaural level difference based on anthropometry and its effect on sound localization,” *The Journal of the Acoustical Society of America.*, 122, pp.2832-2841, 2007
- [65] W. M. Hartmann and Z. A. Constan, “Interaural level differences and level-meter model,” *The Journal of the Acoustical Society of America.*, 112, pp.1037-1045, 2002
- [66] M. Morimoto and H. Aokata, “Localization cues of sound sources in the upper hemisphere,” *J. Acoust. Soc. Jpn. (E)*, 5, pp.165–173,1984.
- [67] M. Morimoto, K. Iida and M. Itoh, “Upper hemisphere sound localization using head-related transfer functions in the median plane and interaural differences,” *Acoust. Sci. & Tech.*, 24(5), pp. 267-275, 2003
- [68] B. F. G. Katz and M. Noisternig, “A comparative study of interaural time delay estimation methods,” *The Journal of the Acoustical Society of America*. 135, pp.3530-3540, 2014

- [69] S. Chatterjee and A. S. Hadi, "Regression Analysis by Example," 5th edition, John Wiley and Sons, Hoboken, NJ, 2012.
- [70] R. S. Woodworth and H. Schlosberg, "Experimental Psychology," Holt, Rinehart and Winston, pp.349–361, 1954
- [71] A. W. Mills "On the minimum audible angle," The Journal of the Acoustical Society of America., 30, pp.237-246, 1958
- [72] 西岡 伸介, 石井 要次, 飯田 一博, "上昇角知覚に関する頭部伝達関数の第 1 ピーク周波数の弁別閾," 日本音響学会講演論文集, pp.859-860, 2013
- [73] K. Iida, and Y. Ishii, "3D sound image control by individualized parametric head-related transfer functions," in Proc. Inter-Noise 2011, Osaka, Japan, 428959, 2011
- [74] J. Blauert, "Sound localization in the median plane," Acustica, 22, pp.205–213, 1969/70
- [75] K. Iida, S. Nishioka and Y. Ishii, "Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener's pinnae," The Journal of the Acoustical Society of America. 136, pp.317–333, 2014
- [76] S. Carlile, P. Leong, and S. Hyams, "The nature and distribution of errors in sound localization by human listeners," Hear. Res. 114, pp.179–196, 1997
- [77] R. Sottek and K. Genuit, "Physical modeling of individual head-related transfer functions," Proc. DAGA, 1999

本論文に関する発表論文

1. 査読付論文

- (1) K. Iida and Y. Ishii, "Effects of adding a spectral peak generated by the second pinna resonance to a parametric model of head-related transfer functions on upper median plane sound localization," *Applied Acoustics*, 129, 239-247, 2018
- (2) 石井 要次, 飯田 一博, "受聴者の頭部形状による両耳間差キューの個人化ー両耳間時間差の推定ー," *日本バーチャルリアリティ学会論文誌*, 22, 3, pp. 405-412, 2017

2. 解説論文

- (1) 石井 要次, 竹本 浩典, 飯田 一博, "耳介形状と頭部伝達関数のなぞ," *日本音響学会誌* 71, pp.127-135, 2015

3. 国際会議

3-1. 口頭発表

- (1) K. Iida and Y. Ishii, "Individualization of the head-related transfer functions on the basis of the spectral cues for sound localization," *Proceedings of IWPASH*, 2009
- (2) Y. Ishii and K. Iida, "Motion-tracked 3D sound image localization system with individualization of head-related transfer functions," *Proceedings of IWPASH*, 2009
- (3) K. Iida, N. Gamoh, Y. Ishii, and M. Morimoto, "Contribution of the early part of the head-related impulse responses to the formation of two spectral notches of vertical localization cues," *Proceedings of FORUM ACUSTICUM*, 2011
- (4) K. Iida and Y. Ishii, "3D sound image control by individualized parametric head-related transfer functions," *Proc. internoise*, 2011
- (5) Y. Ishii and K. Iida, "Estimation of spectral notch frequencies of the individual head-related transfer function from anthropometry of listener's pinna," *The Journal of the Acoustical Society of America*. 133, 3516 / *Proceedings of*

4. 国内の学会等

4-1. 口頭発表

- (1) 石井 要次, 和田 万正, 蒲生 直和, 飯田 一博, "個人に適合した頭部伝達関数の探索方法に関する一考察," 日本音響学会講演論文集, pp. 521-522, 2009
- (2) 石井 要次, 蒲生 直和, 飯田 一博, "スペクトラルキューに基づいた頭部伝達関数の個人化方法とその精度について," 日本音響学会講演論文集, pp. 581-584, 2010
- (3) 飯田 一博, 蒲生 直和, 石井 要次, "頭部伝達関数の第 1 ・第 2 ノッチの検出方法に関する一考察," 日本音響学会講演論文集, pp.473-476, 2010
- (4) 蒲生 直和, 石井 要次, 飯田 一博, "仰角知覚のスペクトラルキューの形成における初期頭部インパルス応答の寄与," 日本音響学会講演論文集, pp. 517-520, 2011
- (5) 石井 要次, 飯田 一博, "正中面内のパラメトリック頭部伝達関数と両耳間時間差による上半球面音像制御," 日本音響学会講演論文集, pp.529-532, 2011
- (6) 石井 要次, 西岡 伸介, 飯田 一博, "正中面のスペクトラルノッチと耳介形状の個人差に関する考察 一定量の個人差情報を備えた頭部伝達関数データベースの構築ー," 日本音響学会講演論文集, pp.463-466, 2012
- (7) 石井 要次, 西岡 伸介, 飯田 一博, "受聴者の耳介形状による頭部伝達関数のスペクトラルノッチ周波数の推定," 日本音響学会講演論文集, pp.593-596, 2013
- (8) 飯田 一博, 石井 要次, 西岡 伸介, "耳介形状から推定したスペクトラルノッチ周波数に基づいた頭部伝達関数の個人化," 日本音響学会聴覚研究会資料, H-2013-91, 2013
- (9) 石井 要次, 西岡 伸介, 飯田 一博, "受聴者の耳介形状に基づいた頭部伝達関数の個人化精度の検証," 日本音響学会講演論文集, pp.855-858, 2013
- (10) 西岡 伸介, 石井 要次, 飯田 一博, "上昇角知覚に関する頭部伝達関数の第 1 ピーク周波数の弁別閾," 日本音響学会講演論文集, pp. 859-860, 2013
- (11) 石井 要次, 海老沼 佑輔, 飯田 一博, "頭部伝達関数のノッチ周波数の個人差と音像定位精度の関係," 日本音響学会講演論文集, pp. 963-964, 2014
- (12) 飯田 一博, 石井 要次, 西岡 伸介, "受聴者の耳介形状による頭部伝達関数の個人化：再考," 電子情報通信学会技術研究報告, 応用音響 114(3), pp.1-11, 2014

- (13) X. Yan, K. Iida, and Y. Ishii, "Comparison in frequencies of spectral peaks and notches and anthropometric of pinnae between HRTF databases," 電子情報通信学会, pp. 43-48, 2014
- (14) 宮本 雄太, 石井 要次, 飯田 一博, "正中面の best-matching 頭部伝達関数と両耳間時間差による 3 次元音像制御," 日本音響学会講演論文集, pp. 605-608, 2014
- (15) 石井 要次, 燕 学智, 飯田 一博, "頭部伝達関数データベースの比較-スペクトラルピークノッチ周波数および耳介形状パラメータの分析-," 日本音響学会講演論文集, pp.615-618, 2014
- (16) 向井 玄典, 石井 要次, 大竹 隼人, 矢野 修平, 飯田 一博, "単一エコーの入射方向が単語理解度に及ぼす影響," 日本音響学会聴覚研究会資料, H-2015-76, 2015
- (17) 飯田 一博, 石井 要次, "Parametric HRTF 再考 -N1・N2・P1+P2 による上方音像制御精度の改善-," 日本音響学会講演論文集, 445-448, 2016
- (18) 石井 要次, 木崎 尚也, 吉田 恵里, 飯田 一博, "受聴者の頭部形状による両耳間時間差の推定 -重回帰モデルの再検討-," 日本音響学会講演論文集, pp.877-880, 2016
- (19) 飯田 一博, 石井 要次, "頭部伝達関数の第 2 ピークが正中面上方の音像定位に及ぼす影響," 電子情報通信学会, 2016
- (20) 石井 要次, 木崎 尚也, 吉田 恵里, 飯田 一博, "受聴者の頭部形状による両耳間時間差の推定," 日本音響学会講演論文集, pp. 447-448, 2015
- (21) 高橋 和夫, 石井 要次, 飯田 一博, "受聴者の頭部形状による両耳間レベル差の推定 -前後左右の非対称性に対応頭部モデルの導入-," 日本音響学会講演論文集, pp. 685-688, 2017

5. 表彰

平成 23 年 (2011 年) 3 月 一般社団法人日本音響学会 学生優秀発表賞 受賞

平成 27 年 (2014 年) 3 月 一般社団法人日本音響学会 粟屋潔学術奨励賞 受賞

謝辞

千葉工業大学の飯田 一博教授には、本研究を進めるにあたり、終始御指導、御鞭撻を頂きました。さらに、研究面だけではなく、私生活面においても数多くの御指導を頂きました。深く感謝致します。

本論文の作成にあたり、数多くの適切な御指導を頂きました千葉工業大学の陶 良教授、今野 将教授、苜木 禎史教授、宮田 高道教授に深く感謝致します。

実験のご協力を頂いた、千葉工業大学飯田研究室の学生の皆様に深く感謝致します。

私が入所後に本論文を作成する機会を与えて頂きました一般財団法人小林理学研究所の山本 貢平所長に深く感謝致します。また、同研究所の土肥 哲也主任研究員には本論文の執筆にあたり、業務面において多くのご配慮、およびご助言を頂きました。深く感謝致します。

また、本学の先輩である一般財団法人小林理学研究所の小林 知尋氏、株式会社ニューズ環境設計の太田 達也氏、一般財団法人防衛施設協会の森 淳一氏、東京都立産業技術研究センターの渡辺 茂幸氏には、実験方法等の御指導を頂きました。また、卒業後も公私にわたり大変お世話になりました。心より感謝致します。

最後に、温かく見守り、就学の機会を与えて頂きました両親、一緒に研究生活を送った同期および後輩の皆様に感謝致します。