

長岡技術科学大学大学院
工学研究科修士論文

題 目

本人耳介モデルを用いた
ダミーヘッドの小型化に関する研究

指導教員

准教授 杉田 泰則

著 者

電気電子情報工学専攻
信号処理応用研究室
18314681 田村 拓久巳

提出期日

令和4年2月10日

ABSTRACT

A Study on Miniaturization of Dummy Head Using of the Person Himself Auricular Model

Author : Takumi TAMURA
Supervisor : Yasunori SUGITA

In recent years, we have seen the development of services in 3D space, such as 3D video technology and the metaverse represented by virtual reality, as well as the remote nature of various interactions due to changes in social conditions. As a result, the need for entertainment that can produce a more realistic feeling has increased. One way to create a sense of presence is through sound image localization technology, which uses headphones or speakers to make sounds perceived in various directions. This technology is expected to be applied in the field of entertainment, such as movies and games, as well as in the field of medical welfare, such as assistive technology for the visually impaired.

One of the methods of sound image localization technology is to use a dummy head. A dummy head is a recording device made of resin that mimics the shape of a human head, and the sound recorded through the dummy head can have a certain degree of realism. However, in terms of sound image localization, it is difficult to say that high localization is achieved. The reason for this is that the shape of the head and auricle is different from the listener's own, although it resembles a human, and this difference is thought to make sound image localization difficult. In addition, the size of the dummy head is the same as that of a human head because it imitates a human head. This makes it difficult to carry around and record while holding it in one's hand.

In this paper, we propose a dummy head system that can be miniaturized by omitting the head shape and reducing the width of the auricle, and that can be used for sound image localization. First, a model of the person himself auricle is created from the scanned data of the auricle part of the listener, and a stereo microphone is created with only the head part omitted. A comparison of the sound image localization experiment with and without the head revealed that the omission of the head part had no effect on the sound image localization. Next, we created a stereo microphone with a reduced auricular width and conducted sound image localization experiments before and after reducing the auricular width, and found a decrease in sound image localization accuracy due to the change in Interaural Time Difference(ITD). After correcting the changed ITD by using the ITD calculated from the database, the sound image localization experiment was conducted

again, and the sound image localization accuracy did not change before and after the reduction, and the auricular width could be reduced.

In order to examine the effectiveness of the proposed method, we compared its localization accuracy with that of actual listening, and found that the proposed method had about 20% lower localization accuracy. It was found that this was caused by a misalignment of the notch in the head transfer function, which indicates the spatial characteristics near the head. This may be attributed to the low accuracy of the person himself auricular model created by the proposed method. As a method of improvement, the shape of the auricle could be scanned by MRI or other precise measurement. In addition, this paper only examines the horizontal plane, and does not examine the median plane. The median plane information is also important to provide a more realistic sound, so this is an issue for the future.

目次

第 1 章	はじめに	1
1.1	研究背景	1
1.2	研究目的	2
1.3	論文構成	2
第 2 章	本研究の基礎となる知識・手法	3
2.1	音像定位技術と座標系	3
2.1.1	音像定位技術	3
2.1.2	座標系	3
2.2	音像定位をさせる主な手法	4
2.2.1	頭部伝達関数を用いた手法	4
2.2.2	ダミーヘッドを用いた手法	5
2.3	ダミーヘッドによる手法の先行研究	5
2.3.1	縮小したダミーヘッドによる音像定位	5
2.3.2	音像定位における耳介形状の重要性	6
2.3.3	ダミーヘッドの簡略化可能形状	6
2.4	両耳間時間差	6
2.4.1	頭部インパルス応答の測定	7
2.4.2	両耳間時間差の算出	7
2.5	両耳間レベル差	9
2.5.1	両耳間レベル差の算出	9
2.6	まとめ	10
第 3 章	提案手法	11
3.1	頭部部分の省略	11
3.1.1	耳介モデルの作成	11
3.1.2	ステレオマイクの作成	12

3.1.3	確認実験	12
3.1.3.1	実験内容	12
3.1.3.2	実験条件	13
3.1.3.3	実験結果・考察	14
3.2	耳介幅の省略	16
3.2.1	耳介幅を縮小したステレオマイク	17
3.2.2	確認実験	17
3.2.2.1	実験内容	17
3.2.2.2	実験条件	17
3.2.2.3	実験結果・考察	18
3.2.3	定位精度低下の原因	20
3.2.3.1	両耳間レベル差	20
3.2.3.2	両耳間時間差	21
3.2.4	両耳間時間差の補正	23
3.2.4.1	補正手法の概要	23
3.2.4.2	補正に使用したデータベース	23
3.2.5	補正手法の詳細	24
3.2.6	補正の確認実験	25
3.3	まとめ	27
第4章	提案手法の検証	28
4.1	被験者の実聴時定位実験	28
4.1.1	実験内容	28
4.1.2	実験条件	28
4.1.3	実験結果	29
4.2	提案手法との比較	31
4.3	定位精度低下の原因	34
4.3.1	帯域制限ホワイトノイズによる定位実験	34
4.3.1.1	実験条件	35
4.3.1.2	実験結果	35
4.3.2	中周波帯域における頭部伝達関数	37
4.4	まとめ	39
第5章	おわりに	40
	謝辞	42

付録 A	他被験者の実験結果	43
A.1	他被験者の帯域制限音像定位実験結果	43
A.2	他被験者のノッチ位置の周波数差	46
参考文献		48

第 1 章

はじめに

1.1 研究背景

近年，バーチャルリアリティ (VR:virtual reality) を代表とする 3 次元映像技術が一般的となり，さらにはメタバースといった 3 次元バーチャル空間でのサービス等も展開されてきている．また，最近の社会的情勢の変化によって通話や会議，仕事までも遠隔でのやり取りになることが多くみられるようになった．これらに伴って，より臨場感を演出することのできるエンターテインメントが求められるようになった．臨場感を演出する方法の一つに，ヘッドホンやスピーカーを用いて様々な方向に音を知覚させる音像定位技術 [1] がある．音像定位技術とはヘッドホン等から再生する受聴音に対し方向感を付与する技術である．この技術は，映画やゲーム，アミューズメント，遠隔でのスポーツ観戦 [2] といったエンターテインメントの分野から，視覚障害者向けの支援技術 [3] といった医療福祉分野，といった幅広い分野へも応用が期待されている．

音像定位技術の手法のもっとも一般的なものとして，頭部伝達関数 (HRTF:Head Related Transfer Function) [1] [4] を用いた方法がある．HRTF とは，ヒトが音の方向を知覚するために手掛かりにしている情報であり，自由空間のある位置から放射された音が音源から鼓膜に届くまでの伝達関数であり，頭部や胴体，特に耳介によって決定される [5]．HRTF は周波数領域に対して使われるのに対し，これをフーリエ変換して時間領域で表現した頭部インパルス応答 (HRIR:Head Related Impulse Response) を再生したい音源に対して畳み込むことによって音像定位をさせることが出来る．高い精度の音像定位を行うためには受聴する本人の HRIR を畳み込む必要がある．高い精度の音像定位を提供できる HRTF を用いた手法だが，複数の問題点が存在する．主だった問題点は二つある．一つ目は HRIR の取得手法である．HRIR は受聴者に対する音源の位置，主に角度によって異なる．そのため，精度を高めるためには多くの角度からの測定が必要であり，その分測定時間がかかることから受聴者への負担も大き

くなる。また、計測の際には音の反射等のない無響室といった特殊な設備で測定を行う必要があり、通常の部屋などでの測定は困難となっている。このため、HRIR を気軽に取得することは困難である [6]。二つ目は、HRIR の個人性についてである。他人の HRIR での音像定位を行う場合、高い精度を得られることは稀有であり、ほとんどの場合は主に前後方向の誤知覚を発生させる [7]。HRIR は頭部や胴体、耳介形状によって大きく決定され、似たような体格の人の HRIR であっても誤知覚を起こすことは多い。そのため、測定のための時間や設備、受聴者の負担を解消するためにダミーヘッドを用いた音像定位手法について研究が行われている。詳しい先行手法・研究については次章に述べる。

1.2 研究目的

HRIR の測定時のような特殊な設備を必要とせず、受聴者の情報収集の負担を低減させ、小型化を行った取り扱いやすいダミーヘッドシステムを提案する。被験者の耳介部スキャンデータから作成した耳介モデルを用い、頭部部分の省略、両耳介幅の縮小を行った小型ダミーヘッドを作成。これを用いて収録したホワイトノイズを用いて被験者の音像定位実験を実施する。その後、被験者の実聴時との音像定位精度を比較することで提案法の有効性を検討する。

1.3 論文構成

本論文の構成を以下に示す。第2章では本論文で用いる基礎的知識及び先行手法について述べる。第3章では、本論文で提案する頭部省略についてと両耳介幅の縮小について、またその時に発生する問題の補正方法について述べる。第4章では提案手法と実聴時で音像定位実験を行い、定位精度の比較を行う。第5章では本論文の結論を述べる。

第 2 章

本研究の基礎となる知識・手法

本章では、本論文で用いる基礎的な知識・先行手法について述べる。2.1 節では音像定位技術と本論文で扱う座標系について述べる。2.2 節では音像定位をさせるための主な二つの手法について、2.3 節ではダミーヘッドによる音像定位に関する先行手法について述べる。また、2.4 節、2.5 節ではヒトが音像定位の手がかりとしている要素である両耳間時間差と両耳間レベル差について述べる。

2.1 音像定位技術と座標系

2.1.1 音像定位技術

音源からの音波が鼓膜に到達した際、ヒトがその音の方向や距離を知覚する時、この音を音像という。また、この音像の位置を知覚する能力を音像定位能という。さらに、ヘッドホンなどから音を受聴する際に再生音に対して方向や距離等の情報を付与し、受聴者に再生音を音像として知覚させる技術のことを音像定位技術という。ヘッドホンでの高い音像定位技術の実現には大きく二つの手法が存在する。一つ目は頭部伝達関数を用いた手法、もう一つがダミーヘッドによる録音を用いた手法である。本論文では後者のダミーヘッドによる手法をベースにするが、本論文の検討にも必要となるため両手法について次節で述べる。

2.1.2 座標系

ここでは、本論文で音源の方向を表す際の座標系について述べる。本論文では図 2.1 に示している、受聴者頭部を中心とした球面座標系を用いる。この座標系の定義は次の通りである。座標原点は左右耳の外耳道入り口を結ぶ線分の中点とする。水平面は眼窩点(眼球の入る頭蓋穴の下縁の最も低い点)と耳珠(外耳道入り口に存在する突出

している部位)を含んだ平面であり、横断面は左右外耳道入り口を通過する水平面に直交する面である。また、正中面は水平面、横断面両方に直交する面となっている。角度表記は方位角を θ 、仰角を φ とする。

なお、本論文では音源が水平面上に存在している場合に対する実験、検証を行っているため仰角 $\varphi=0^\circ$ となるため、これ以降 φ は省略するものとする。水平面に関しては、図 2.1 に示す正面を $\theta=0^\circ$ とし、時計回りに $0^\circ\sim180^\circ$ と増加していくものとする。

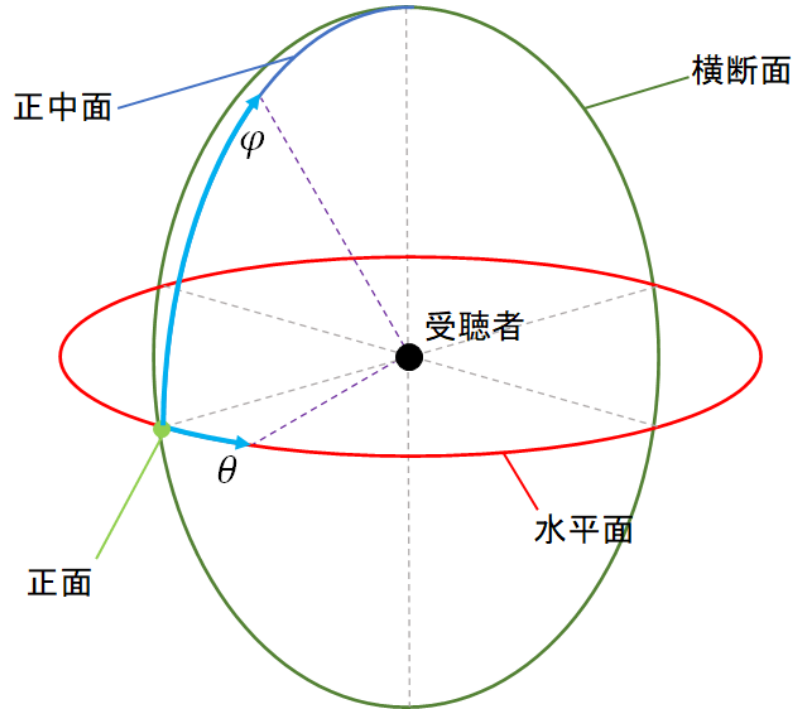


図 2.1: 球面座標系

2.2 音像定位をさせる主な手法

2.2.1 頭部伝達関数を用いた手法

ある音波が鼓膜に到達する際に、受聴者の頭部や胴体、耳介の形状の影響を受ける。このような頭部周辺によって生じる音の特性変化を周波数領域の関数で表したものを頭部伝達関数 (HRTF: Head Related Transfer Function) という。また、時間領域で表したものは頭部インパルス応答 (HRIR: Head Related Impulse Response) という。頭部伝達関数は次式のように表される。

$$H_{L|R}(\theta, \omega) = \frac{G_{L|R}(\theta, \omega)}{F(\theta, \omega)} \quad (2.1)$$

ここで $H_{L|R}(\theta, \omega)$ は左右の HRTF であり, θ は音源の方位角を, ω は周波数を表す. $G_{L|R}(\theta, \omega)$ は, 反射, 回折の無い自由音場における受聴者の左右外耳道入り口までの伝達関数である. $F(\theta, \omega)$ は, 自由音場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.

頭部伝達関数によって音像定位をさせるには, 再生音に対して HRIR を畳み込むことによって再現することができる. このように非常に簡単な処理で音像定位をさせることが出来る HRTF だが, 前章でも述べたように, 精度の高い定位には本人の HRTF が必要不可欠であり, HRTF の測定には無響室などの設備や空間の伝達関数を測定するため受聴者は測定中に動くことが出来ない. また, 角度に依存した情報であるため必要な角度の分だけ測定が必要となる. このため受聴者に負担もかかり, 気軽に扱うことは難しいと言える.

2.2.2 ダミーヘッドを用いた手法

ダミーヘッドとは, 樹脂などでつくられた人間の頭部を模した録音機器のことである. ダミーヘッドを用いて録音した音はある程度の臨場感を得ることは可能である. しかしながら音像定位といった面では精度が高い定位を得られるとは言い難い. 理由としては, 前節で述べた HRTF が原因として挙げられる. 頭部や耳介形状を模しているとはいえ, 本人のもので無いとやはり高い定位を得られない.

また, それ以外の問題点も存在する. ダミーヘッドはヒトの頭部を模しているためサイズとしては人の頭部と同様の大きさをしている. このため, 持ち運ぶ際に手間となったり, どこかに固定して置く際にも安定性の高くしっかりとした土台が必要である. また, 手に持ちながら録音するといった用途にも対応しづらい. しかしながら, 直接録音をするため HRTF のような計算時間を必要とせず, リアルタイム性は非常に高い. また, 受聴者の頭部を模したダミーヘッドであれば HRTF を模していることになるため, 高い音像定位を得ることは可能である.

2.3 ダミーヘッドによる手法の先行研究

2.3.1 縮小したダミーヘッドによる音像定位

平原ら [8] はダミーヘッドの持ち運び等の不便性の解決のために受聴者本人の頭部, 耳介形状をスキャンし, そのサイズを 100%, 70%, 50% と縮小して頭部モデルを作成, それらを用いて無響室でホワイトノイズを収録, 被験者にヘッドホンを用いて聞いてもらい音像定位実験を行った.

結果としては, 100% サイズの頭部モデルでは非常に高い定位精度を得ることが出

来ていたが、70%、50%と縮小するにつれ定位精度は著しく低下してしまった。

原因としては頭部や耳介部分の縮小に伴って、HRTFが複雑に伸張してしまい、本人の頭部や耳介形状を模していても音像定位をさせることはできなかったと考えられる。このことから、小型化の際に頭部、耳介部の縮小は定位に大きく影響を及ぼすことが分かる。

2.3.2 音像定位における耳介形状の重要性

中村ら [9] は、高い精度の音像定位には耳介形状の情報が重要であり、耳介を模しているのであれば頭部の形状の本人性は必要ないと考え、ダミーヘッドに対して被験者本人の耳介モデル、被験者でない他人の耳介モデルを用意し、それぞれの耳介で音像定位実験を行った。

結果として、他人の耳介モデルを用いた際には定位が出来なかったが、本人の耳介モデルを用いた際にはしっかりとした定位を出すことが出来ていた。このことから、頭部形状は平均的なモデルを用いても問題ないが、耳介形状は本人のものをを用いる必要性があることが判明した。

2.3.3 ダミーヘッドの簡略化可能形状

戸嶋ら [10] は、被験者の頭部モデル作成の際にどの程度までの簡略化が可能であるかの実験として、耳介形状以外の部分をラグビーボール状の球体にまで簡略化を行い、頭部形状まで再現したダミーヘッドと、簡略化したダミーヘッドとの定位精度比較の実験、また、被験者本人の耳介モデルと他人の耳介モデルを用いた比較の実験を実施した。

結果として、頭部形状の再現の有無に関しては大きな音像定位精度の変化は見受けられず、頭部形状がヒト頭部の平均的な形でなくとも耳介形状が本人モデルであれば球体でよいという結果に、また、耳介モデルは本人のものでないと前後誤りが発生してしまい、正確に定位が出来ないという結果となった。このことから、頭部形状はヒト頭部のような鼻部の突起等は必要なく、単純な球体のような形状でよいことと、耳介モデルは本人のものでないと定位が出来なくなることが判明した。

2.4 両耳間時間差

両耳間時間差 (ITD: Interaural Time Difference) とは、ある音源から発せられた音が左右の耳に届く時の時間差のことで、低周波数 (0~2000Hz) の音ではこの ITD がヒトが方向を知覚するのに手がかりとしている成分であることが知られている。これ

は、低周波数の音が高周波数の音に比べて回折しやすいことに起因している。

2.4.1 頭部インパルス応答の測定

ITD の算出には HRIR から算出を行う。本論文における HRIR の測定方向は水平面の 12 方向 (30°) 間隔、スピーカから頭部中心までは 1.2 m である。測定には TSP 信号を用いた。なお、本論文における測定信号及び録音のサンプリング周波数は 48 kHz となっている。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外耳道入り口 (外耳道閉塞状態 [11]) までのインパルス応答 $g_l(t)$, $g_r(t)$ を測定する (2^{16} サンプル)。
2. 受聴者の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インパルス応答 $f(t)$ を測定する (2^{16} サンプル)。
3. 正面方向 (0°) の $g_l(t)$, $g_r(t)$ から振幅の絶対値が最大となる時刻をそれぞれ求め、時刻の早い方から 50 サンプル前を切り出して開始位置 T_0 とする。
4. 全方向の $g_{l,r}(t)$ 及び $f(t)$ を T_0 から 512 サンプルの矩形窓で切り出す。
5. $g_{l,r}(t)$ 及び $f(t)$ をフーリエ変換し、512 サンプルの複素数 $G_{l,r}(\omega)$, $F(\omega)$ を求める。
6. 式 (2.2) から $HRTF(\omega)$ を算出する。

$$HRTF_{l|r}(\omega) = \frac{G_{l|r}(\omega)}{F(\omega)} \quad (2.2)$$

7. $HRTF(\omega)$ を逆フーリエ変換し、HRIR を取得する。

2.4.2 両耳間時間差の算出

2.4.1 小節で求めた HRIR を用いて ITD を算出する。本節のはじめに述べたように、ITD が支配的な周波数帯域は 2000Hz 以下となる。また、ITD には周波数依存性があるために HRIR の両耳間相互相関によって ITD を定義することとした。ITD の算出方法は以下の通りである [12] [13]。また、計算時に用いるローパスフィルタは窓関数法によって設計を行い、設計仕様を表 2.1 に、図 2.2, 図 2.3 に作成したフィルタの振幅特性、位相特性を示す。なお、特性グラフ中の縦の赤破線は 2000Hz を示している。

1. HRIR に対し 2000Hz を通過域端周波数とする FIR ローパスフィルタを畳み込む。
2. 時間分解能を上げるため、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 (2.3)) が最大となる時間差 τ を ITD とする [12].

$$\Phi_{l,r}(\tau) = \lim_{T \rightarrow \infty} \frac{\int_{-T}^T HRIR_l(t) \times HRIR_r(t - \tau) dt}{\sqrt{\int_{-T}^T HRIR_l^2(t) \times HRIR_r^2(t) dt}} \quad (2.3)$$

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

表 2.1: フィルタ設計仕様

カットオフ周波数	2000Hz
使用窓関数	ハニング窓
フィルタタップ数	255

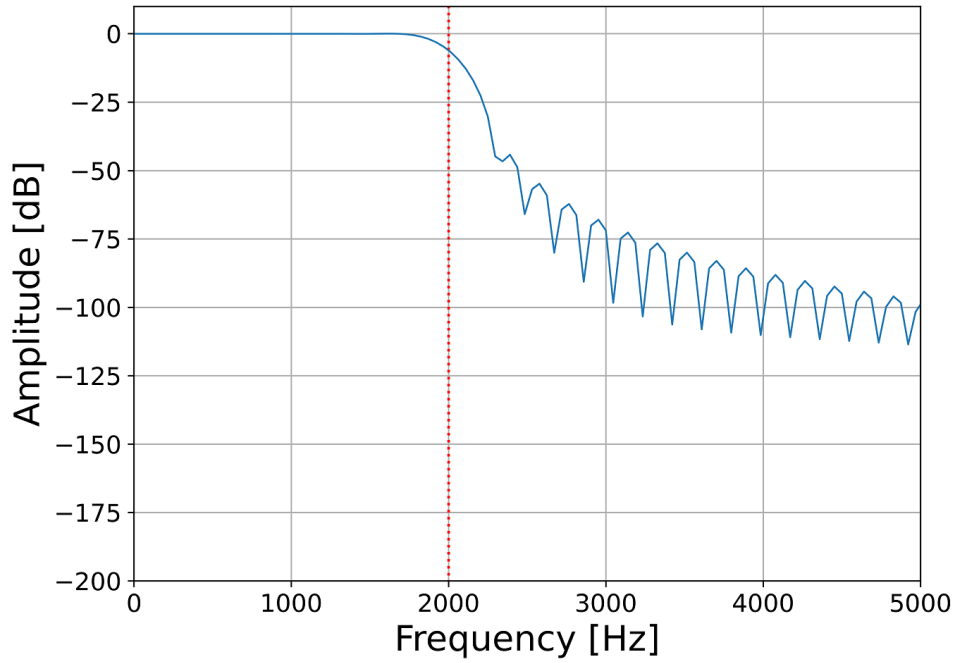


図 2.2: 振幅特性

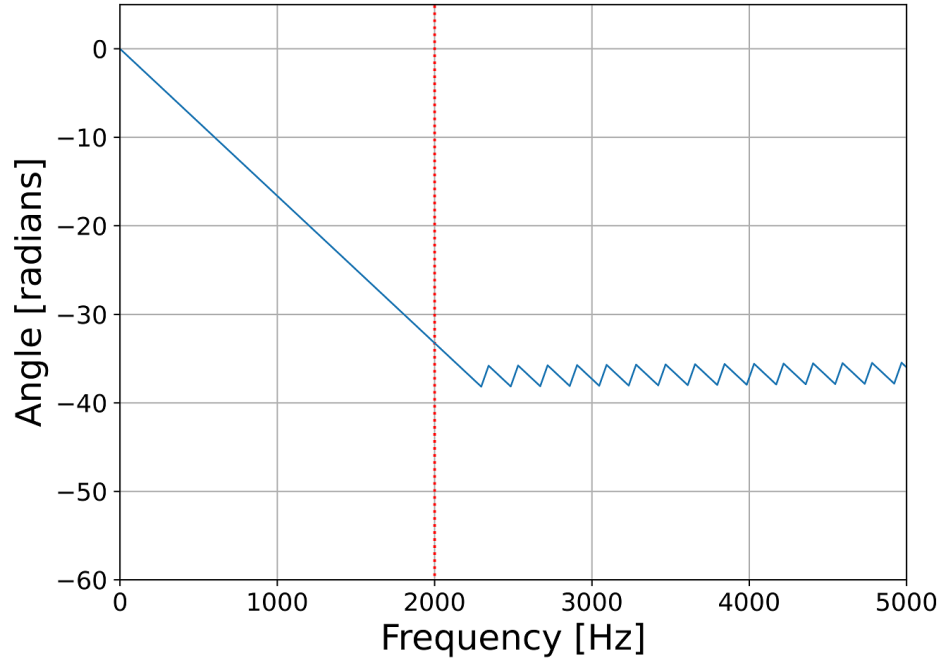


図 2.3: 位相特性

2.5 両耳間レベル差

両耳間レベル差 (ILD: Interaural Level Difference) とは, ある音源から発せられた音が左右の耳に届く時のレベル差のことで, 基本的にすべての周波数でヒトの方向知覚の手掛かりに影響しているが, 高周波帯域になるにつれて影響は大きくなる.

2.5.1 両耳間レベル差の算出

両耳間レベル差の算出についても HRIR を用いて計算を行うため, 2.4.1 小節で求めた HRIR を用いて ILD を算出していく. ILD の算出方向は以下の通りである [14].

1. 512 サンプルの HRIR の後ろに 0 を付加し, 48000 サンプルで FFT を行う (周波数分解能 1Hz).
2. FFT で得られた HRTF のエネルギー E_l , E_r を算出する.
3. 左右の耳のエネルギー E_l , E_r の差を ILD とする (式 (2.4)).

$$ILD = 10 \log_{10}(E_r/E_l) \quad (2.4)$$

2.6 まとめ

本章では本研究の基礎となる知識・手法について述べた。音像定位をさせることが出来る手法として、HRTFを用いた手法とダミーヘッドを用いた手法があり、本論文ではダミーヘッドを用いた手法について述べていく。ダミーヘッドを用いた手法については特殊な測定設備を必要としない反面、頭部や耳介の形状が本人のものでないと高い定位は得られない。また、ヒト頭部を模した録音機器であるためにサイズが大きく、持ち運びづらい、設置する際にも丈夫なものが必要、手に持ちながらといった用途に対応しにくいといった実用時の問題点がある。これらの問題点の解決のため、全体を縮小したダミーヘッドによる音像定位実験 [8] が行われたが、頭部や耳介が小さくなったことで HRTF が複雑に伸張してしまった。文献 [9]、文献 [10] では定位には耳介の影響が大きいとして簡易的、または本人とは異なった形状の頭部に本人耳介モデルを取り付けることによって高い定位精度を出すことが出来ることが示された。3章では、音像定位をさせることが出来るダミーヘッドの小型化の提案手法として、本人耳介モデルを用い、両耳介幅の縮小、頭部の省略をしたダミーヘッドの概要と縮小・省略による影響の有無の確認実験、補正手法と補正の確認実験について述べる。

第 3 章

提案手法

2.3 節では、ダミーヘッドによる先行手法について述べた。先行手法より、頭部部分を省略したとしても耳介形状さえ本人のものであれば定位が出来る可能性は十分に考えられる。また、頭部部分の省略によって耳介幅の省略も可能であると考えられる。

本章では、提案手法について説明する。3D スキャナ、3D プリンタによって作成した耳介モデルを、頭部部分を省略したステレオマイク、耳介幅を縮小したステレオマイクそれぞれに装着、被験者による音像定位実験を行い、それぞれの有効性を確認する。3.1 節では、耳介モデルの作成、頭部部分の省略について示す。3.2 節では、耳介幅の縮小について示す。

3.1 頭部部分の省略

本節では、頭部部分を省略したステレオマイクの作成についてと確認実験について示す。まず、被験者の耳介をスキャンし、耳介モデルの作成をする。次に、耳介幅を頭部と同様の幅にし、頭部部分を省略したステレオマイクを作成、耳介モデルと組み合わせる。作成したステレオマイクとダミーヘッドに対して耳介モデルを取り付けたものでホワイトノイズを収録し、その音声を用いて音像定位実験を行い、頭部部分の省略の有効性について確認する。

3.1.1 耳介モデルの作成

被験者耳介のスキャンにはハンディタイプの 3D スキャナ (Artec Eva) を用いる。被験者には椅子に座ってもらい、5～10 分程度静止してもらう。その間に耳介のスキャンを行う。5～10 分の被験者の静止拘束時間があり、被験者の負担があるように考えられる。しかし、HRTF の測定時には測定の分解能にもよるが、30° 刻みの測定でも約 30 分ほどの静止拘束時間がある。それと比較すれば、提案手法の拘束時間は短いと言

える。

スキャンした耳介のデータを用いて 3D プリンタ (Flashforge GuiderII) を用いてシリコン型を作成し、シリコンを押し固めて耳介モデルを作成する。

3.1.2 ステレオマイクの作成

3DCAD にて耳介モデル，使用するマイク (SONY ECM-C10) が装着可能で，両耳介の外耳道入り口の幅をダミーヘッドマイク (SAMAR Type4500) と同様である 146.5mm になるものを作成．それらを組み合わせて図 3.1 に示すステレオマイクにした．次節は作成したステレオマイクを用いて頭部部分の省略の有効性の確認実験を行う．

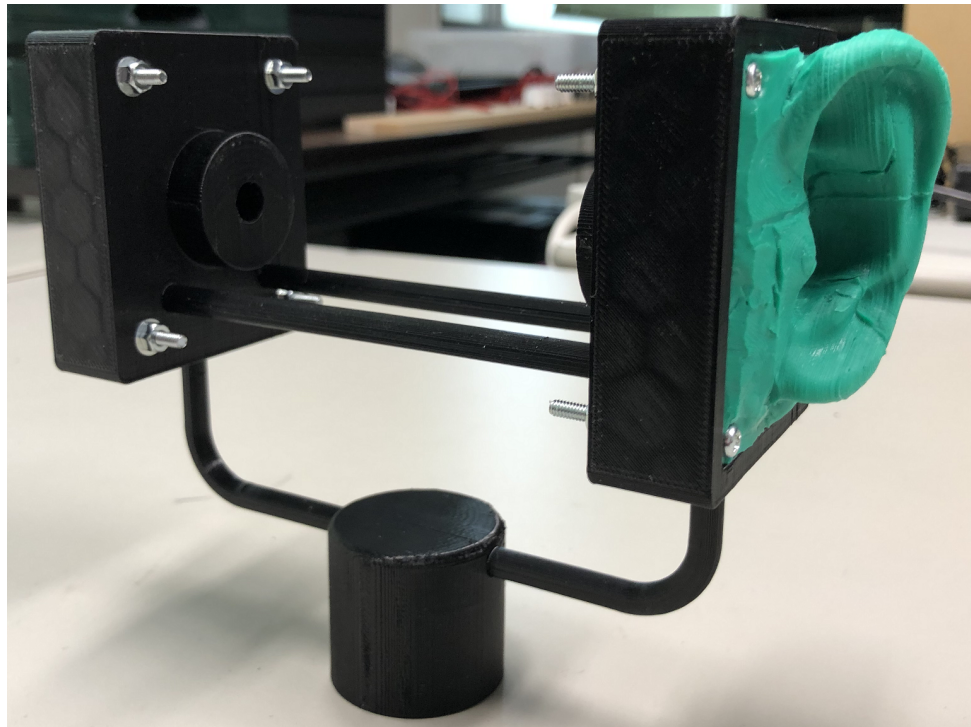


図 3.1: 作成したステレオマイク

3.1.3 確認実験

3.1.3.1 実験内容

頭部部分の省略によって音像定位精度が低下しないことを確認する．具体的な実験内容としては，作成したステレオマイクとダミーヘッドマイク (SAMAR Type4500) に同様の被験者耳介モデルを取り付けた状態で無響室にてホワイトノイズを収録，収録した音声を用いて被験者にヘッドホンを用いて音像定位実験を行ってもらう．実験結

果から頭部部分の省略の有効性を確認する。

3.1.3.2 実験条件

無響室での収録時の条件は表 3.1 のようになっている。また、収録環境の概略図を図 3.2 に示す。図のように、ダミーヘッドマイク、ステレオマイクを囲むように 12 方向 30° 間隔でスピーカーを配置し、ホワイトノイズを収録する。なお、ダミーヘッドマイク、ステレオマイクはスピーカーの中心と外耳道入り口が水平になるよう設置した。また、音像定位実験は表 3.2 に示した条件で行った。被験者はヘッドホンから提示される音を聞き、各方向 5 回の計 60 回の回答を行う。なお、被験者には必ず回答を行ってもらう方式とした。

表 3.1: 収録環境

収録場所	無響室
収録音	ホワイトノイズ
サンプリング周波数	48kHz
角度分解能	30°
音源距離	1.2m
マイクロホン	SONY ECM-C10
スピーカー	MarkAudio PLUVIA7HD-GR-2

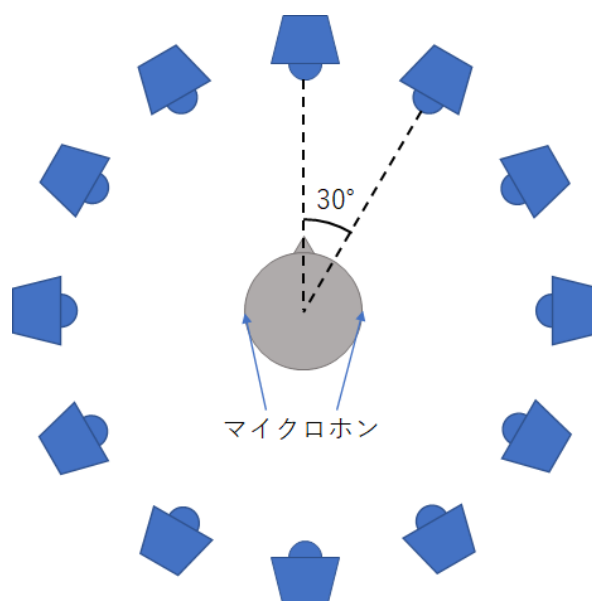


図 3.2: 収録図

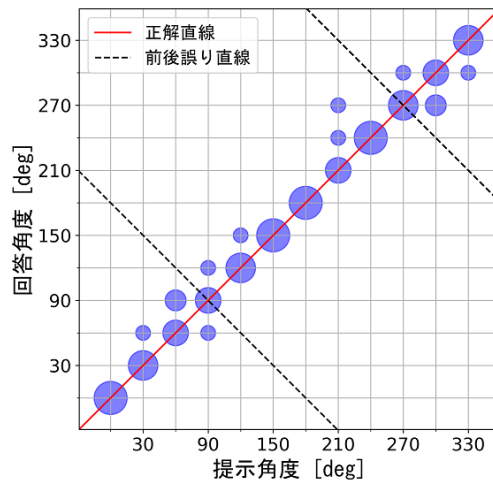
表 3.2: 実験条件

被験者	健聴者 3 名
提示音	ホワイトノイズ
提示回数	60 回 (12 方向 ×5 回)
提示音時間	3sec
ヘッドホン	SONY MDR-CD900ST

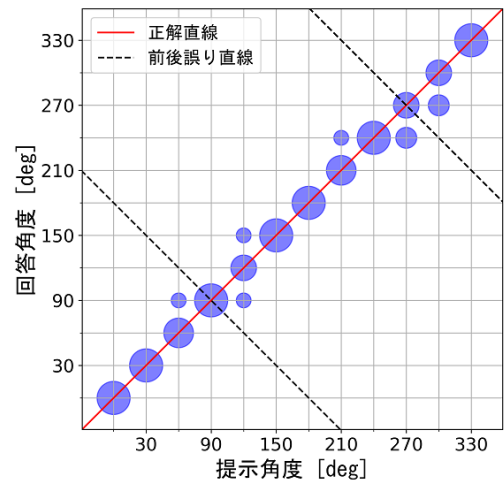
3.1.3.3 実験結果・考察

図 3.3, 図 3.4, 図 3.5 に定位実験の結果を示す. 画像左側グラフはダミーヘッドマイクに耳介モデルを取り付けた際の結果を, 右側グラフは作成した頭部部分を省略したステレオマイクでの定位実験結果を示している. 結果のグラフはバブルチャートで表しており, 横軸が提示音の角度を, 縦軸が提示角度に対する被験者の回答角度を示している. グラフ内にある赤の直線に円がある場合には提示角度に対して被験者が正確に回答していることを示しており, 黒の破線に円がある場合には被験者が提示音に対して前後を誤って認識してしまっていることを示す. また, 円の大きさは回答回数を示している. 各被験者の結果を見ると, 頭部部分の有無にかかわらず同様の定位精度になっているように見受けられ, 前後誤判定もないように見られる. 表 3.3 に赤の直線上に回答のある正確回答の割合を示す. この結果からも頭部部分の省略によって音像定位精度が低下することは無く, ダミーヘッドマイクの小型化の上で頭部部分の省略は有効であると言える.

次節では, 頭部部分を省略したことによって可能となった両耳介幅の省略について実験を行っていく.

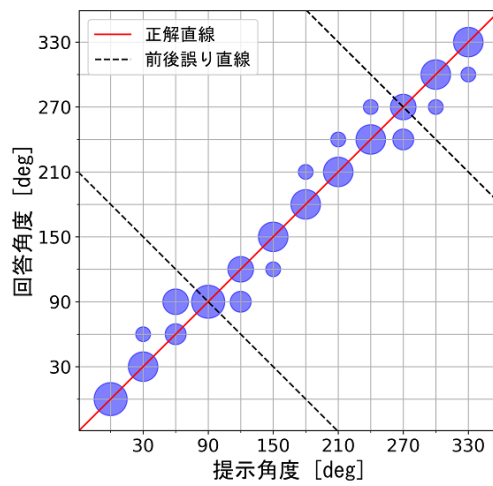


(a) 頭部部分有り

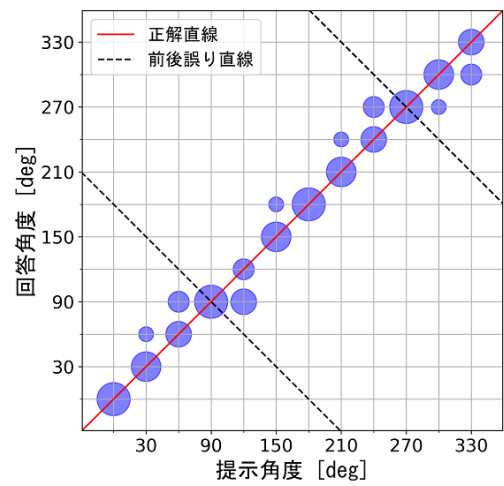


(b) 頭部部分無し

図 3.3: 被験者 A の実験結果



(a) 頭部部分有り



(b) 頭部部分無し

図 3.4: 被験者 B の実験結果

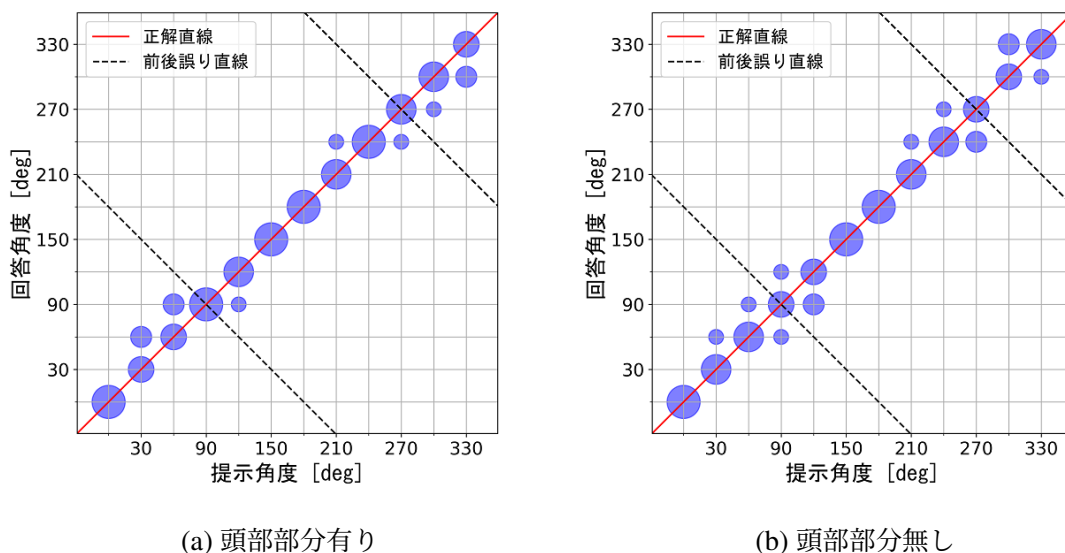


図 3.5: 被験者 C の実験結果

表 3.3: 正確回答割合

被験者	頭部部分無し	頭部部分有り
被験者 A	80%	86.7%
被験者 B	76.7%	78.3%
被験者 C	78.3%	78.3%

3.2 耳介幅の省略

3.1 節では、頭部部分の省略による定位精度に変化はなく、ダミーヘッドマイクの小型化の際に頭部部分は省略可能であることを示した。本節では、頭部部分の省略に伴って両耳介間に空間が出来ることによる耳介幅の省略が可能であると考え、両耳介幅の省略と確認実験について述べる。

耳介モデルは前節で作成したものをを用い、耳介モデル、マイクを入れるステレオマイク部分を両耳介幅が最小となるように設計し、前節と同様の定位実験を行って有効性を確認する。

3.2.1 耳介幅を縮小したステレオマイク

作成する条件は 3.1.2 節と同様であり，耳介幅のみ最小になるよう設計を行って作成した．実際に耳介部分を装着したものを図 3.6 に示す．耳介幅は 49.0mm となった．これは，使用しているマイクの全長が 20mm あること，耳介モデルとステレオマイクの接合部に 10mm ほど必要なことに起因しており，現在の構成だとこの幅が最小となっている．

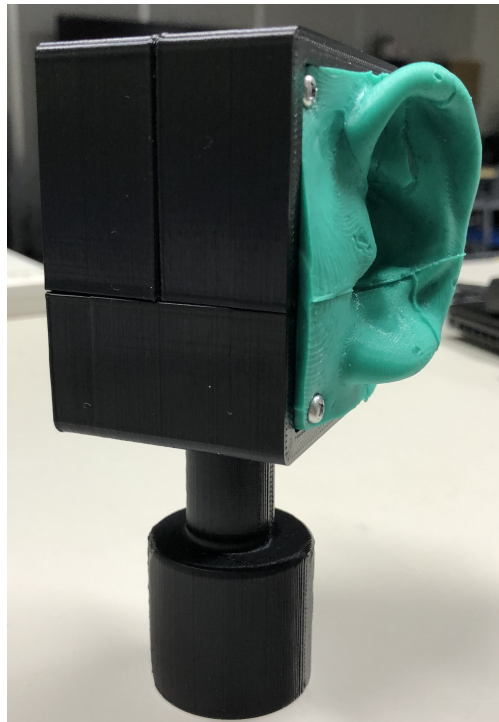


図 3.6: 耳介幅を縮小したステレオマイク

3.2.2 確認実験

3.2.2.1 実験内容

3.1.3.1 節で示したものと同様の実験を行う．実験結果を 3.1 節で作成したステレオマイクでの定位試験結果と比較し，検討を行う．

3.2.2.2 実験条件

実験条件も前節と同様，無響室で収録した 30° 間隔 12 方向のホワイトノイズを被験者にヘッドホンで各方向 5 回の計 60 回をランダムで聞かせる音像定位実験を実施する．

3.2.2.3 実験結果・考察

図 3.7, 図 3.8, 図 3.9 に定位実験の結果を示す. 画像左側は 3.1 節で作成した頭部無しステレオマイクでの定位実験結果を, 画像右側は耳介幅を縮小したステレオマイクでの定位実験結果を示している. 実験結果を見ると, 耳介幅を縮小したステレオマイクでの定位は全被験者で前後の誤判定を発生させており, 音像定位が出来ていないと考えられる. 表 3.4 を見ても耳介幅を縮小した後は著しく音像定位精度が低下していることが分かる. 以上のことから耳介幅を単に縮小すると本人の耳介モデルを用いても音像定位が出来なくなってしまうことが分かった. 次節では定位精度が低下した原因について調査していく.

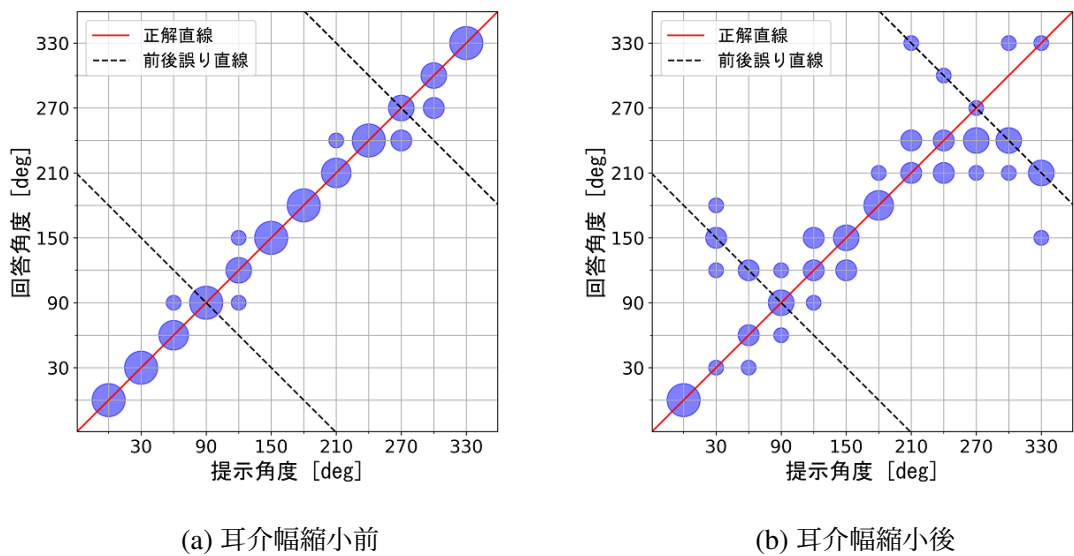
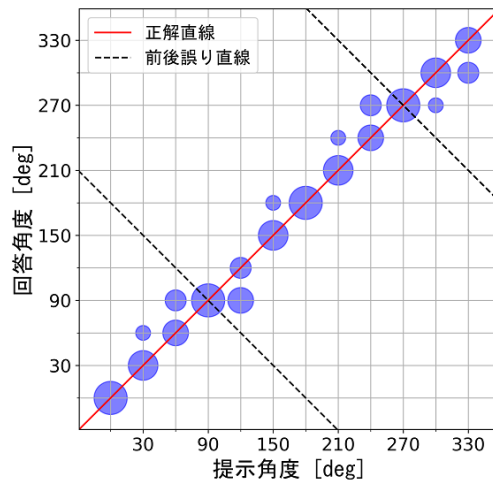
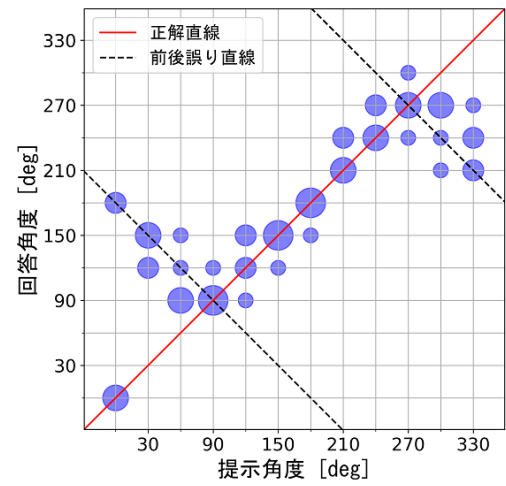


図 3.7: 被験者 A の実験結果

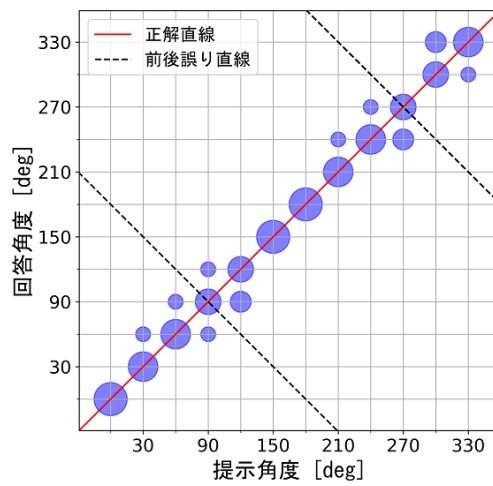


(a) 耳介幅縮小前

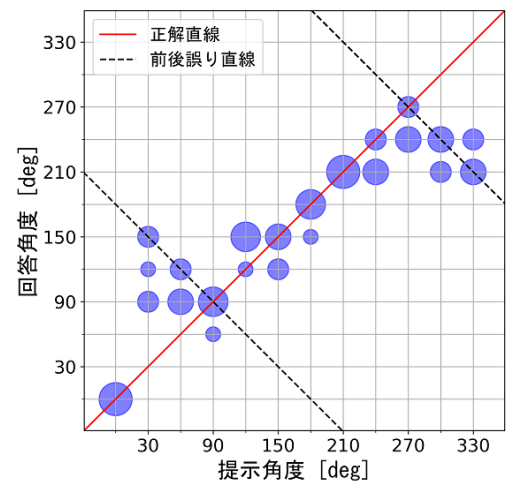


(b) 耳介幅縮小後

図 3.8: 被験者 B の実験結果



(a) 耳介幅縮小前



(b) 耳介幅縮小後

図 3.9: 被験者 C の実験結果

表 3.4: 正確回答割合

被験者	耳介幅縮小前	耳介幅縮小後
被験者 A	80%	43.3%
被験者 B	76.7%	41.6%
被験者 C	78.3%	43.3%

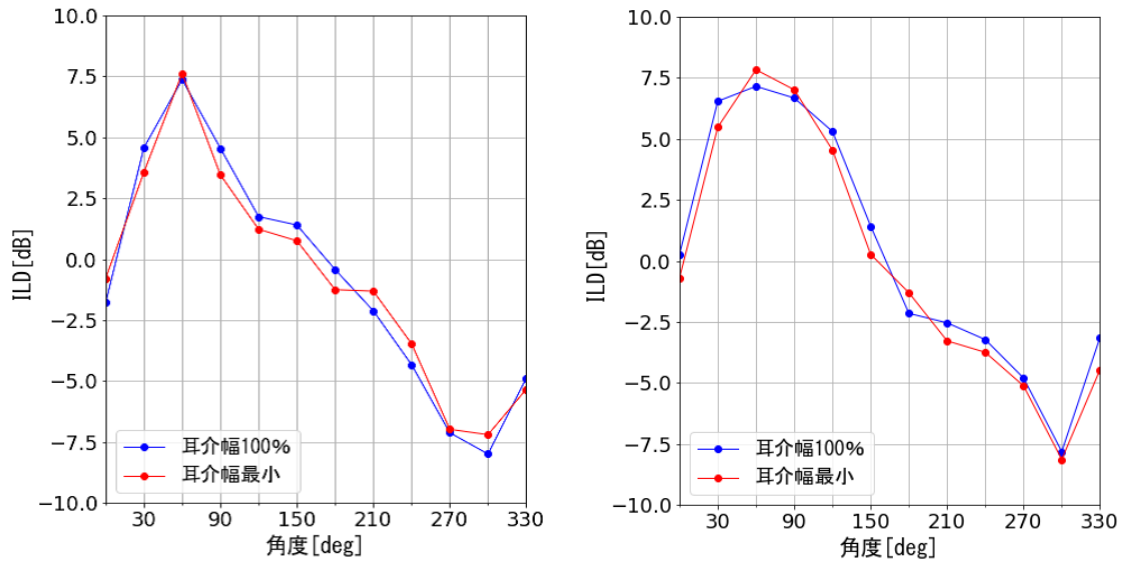
3.2.3 定位精度低下の原因

3.2.1 節では、耳介幅を縮小したステレオマイクでの定位実験を行った。結果としては、耳介幅を縮小した後は前後誤判定が発生しており、音像定位が出来ていないことが分かった。本節では音像定位精度低下の原因として考えられる両耳間レベル差、両耳間時間差について実際に調査を行う。

3.2.3.1 両耳間レベル差

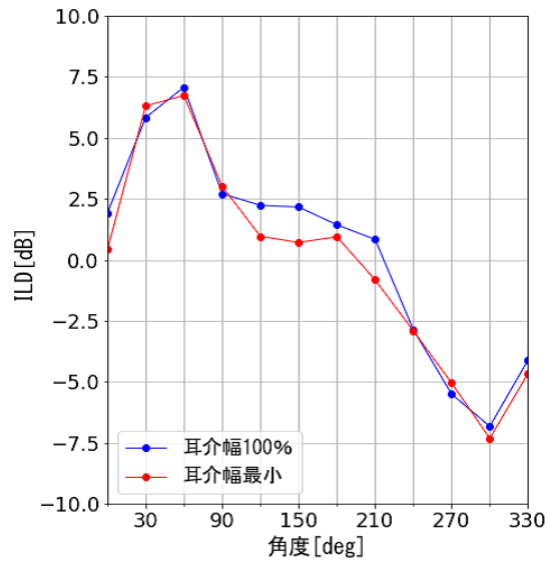
2.5 節で説明したように、両耳間レベル差 (ILD: Interaural Level Difference) とは、ある音源から発せられた音が左右の耳に届く時のレベル差であり、ほぼすべての周波数に作用するが、高周波帯域に大きく作用する。

算出した各被験者の ILD を図 3.10(a), 図 3.10(b), 図 3.10(c) に示す。各被験者の結果を見ると、被験者によってレベル差には個人差があるように見受けられるが、耳介の幅によって大きく差があるようには見受けられず、ILD は定位精度低下の原因としては考えにくい。



(a) 被験者 A の ILD

(b) 被験者 B の ILD



(c) 被験者 C の ILD

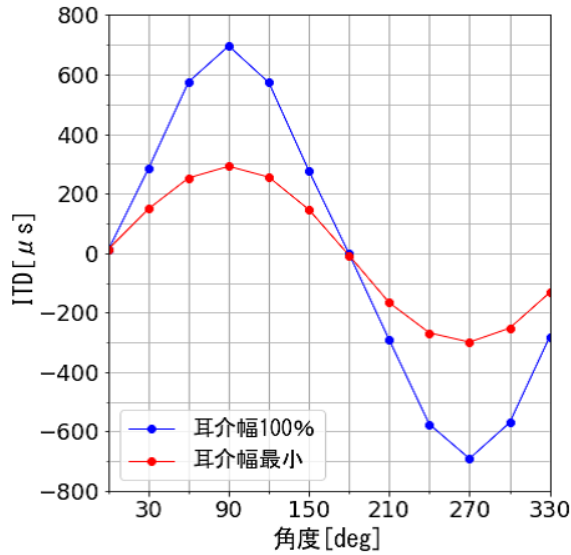
図 3.10: 各被験者の ILD

3.2.3.2 両耳間時間差

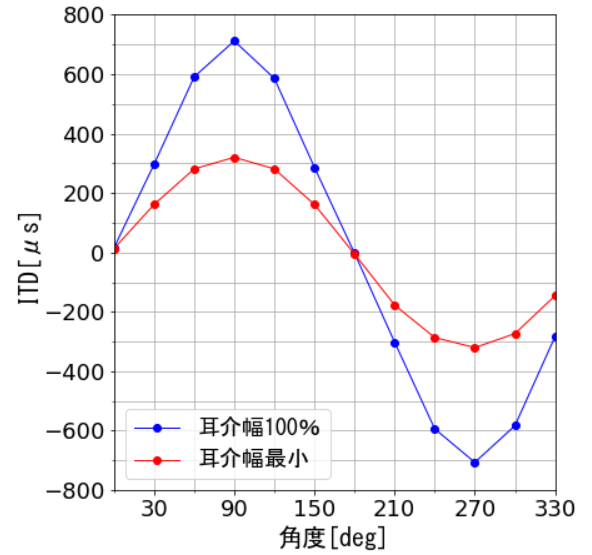
先と同様、2.4 節で述べたように、両耳間時間差 (ITD: Interaural Time Difference) とは、ある音源から発せられた音が左右の耳に届くまでの時間差のことであり、低周波数帯域 ($\sim 2000\text{Hz}$) において大きく作用する。

算出した各被験者の ITD を図 3.11(a), 図 3.11(b), 図 3.11(c) に示す。全被験者 ITD に個人差はさほど見られないが、耳介幅によって大きく変化しており、最大で約 $400\mu\text{s}$

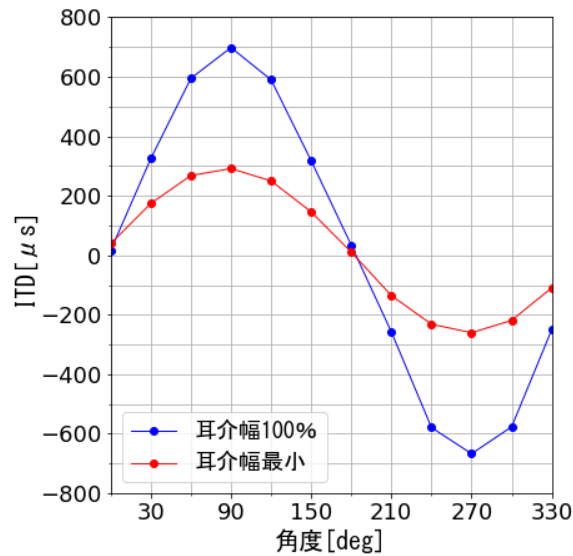
の差があることが分かる。Klump らによれば、ヒトの脳神経回路における時間分解能は約 $19\mu\text{s}$ であると言われている [15] ため、定位精度低下の原因は ITD にあると考えられる。このため、ITD の補正を行い、音像定位可能になるか確認を行う必要がある。次節では、両耳間時間差の補正に関して述べていく。



(a) 被験者 A の ITD



(b) 被験者 B の ITD



(c) 被験者 C の ITD

図 3.11: 各被験者の ITD

3.2.4 両耳間時間差の補正

3.2.3.2 節で示したように，耳介幅の縮小による音像定位精度低下の原因は ITD である可能性が大きいことが分かった．本節ではその ITD の補正を行い，音像定位精度が向上するかの実験を行う．まず，補正の手法について述べ，次に補正に用いるデータについて示す．そして最後に補正後の音源を用いて補正ができたかどうかの確認実験を行う．

3.2.4.1 補正手法の概要

両耳間時間差の補正だが，一番確実な手法としては耳介幅縮小後の ITD を縮小前の ITD に戻す方法であるが，2.4 節で述べたように，ITD の算出には HRIR を用いて算出する必要がある．よって本人の ITD を用いて補正を行うということは，本人の HRIR を用いることと同様であり，2.2.1 節でも述べたように HRIR の取得には無響室などの特殊な環境や被験者の静止拘束時間が長いいため使用したくない．なにより，HRIR を取得できるのであれば HRTF を音源に畳み込む手法を用いることも可能となるため本手法は意味をなさなくなってしまう．そのため，被験者本人の HRIR 情報を用いずに補正を行う必要がある．

3.2.3.2 節の各被験者の ITD を見ると，ITD にほとんど個人差は無いように見受けられる．そこで，データベースにある HRIR の情報を用いて ITD の補正を試みる．次小節に使用するデータベースについて示す．

3.2.4.2 補正に使用したデータベース

データベースの補正には東北大学電気通信大学先端音響情報システム研究室データベース (<http://www.riec.tohoku.ac.jp/pub/hrtf/index.html>) [16] のものを用いる．データベースの概要について表 3.5 に示す．このデータベースの全被験者の ITD を図 3.12(a) に，その平均を取ったものを図 3.12(b) に示す．105 名分のデータを見ても HRIR に個人差はあるが ITD には大きく個人性は無く，おおよそどのヒトでも同様の ITD となると考えられる．このデータベースでの平均値で実際に補正を行い，確認実験を行う．

表 3.5: データベース概要

データベース被験者数	105 名 (内 2 データはダミーヘッドデータ)
使用方向	水平面 0°～330° 30° 刻み
サンプリング周波数	48kHz

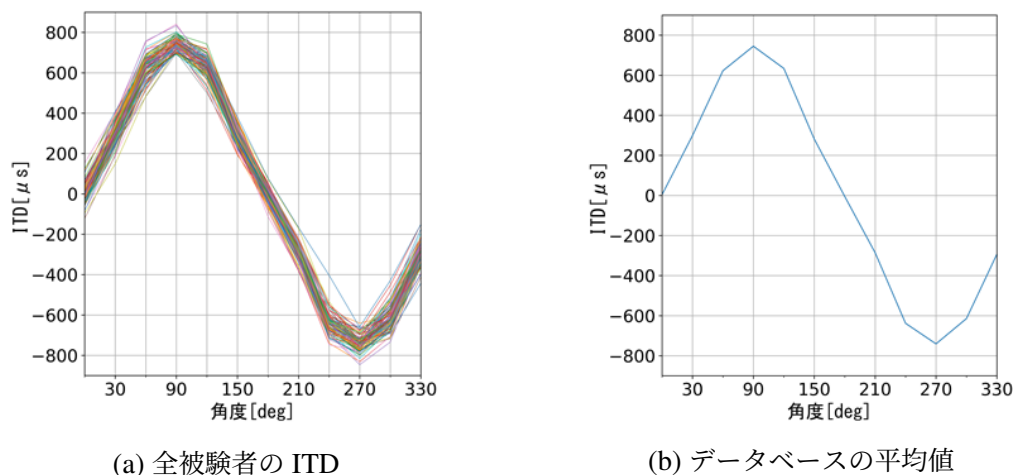


図 3.12: データベースから算出した ITD

3.2.5 補正手法の詳細

前節では、ITD の補正に用いるデータベースについて述べた．ここでは、実際の補正手法について述べる．図 3.13 に示している、図 3.11 に示している耳介幅が最小のデータ 3 名分の平均値と図 3.12(b) に示しているデータベースの平均値を用いて補正する．具体的には、青線と赤線との差を計算し、 $30^{\circ} \sim 150^{\circ}$ は左耳の音源の、 $210^{\circ} \sim 330^{\circ}$ は右耳の音源の開始を差分だけ遅れさせることによって ITD をデータベースと同様の秒数にした． 0° 、 180° に関しては軽微な差があり、どちらも正の方向 (左側が遅れている) に ITD があるため、右側の音を遅らせることによって補正した．

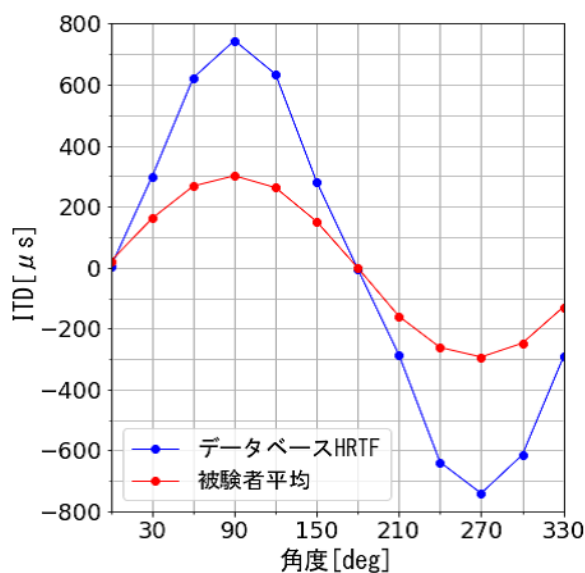


図 3.13: 補正に用いるデータ

3.2.6 補正の確認実験

補正した音源を用いて再度、音像定位実験を行う。実験条件としては先の実験と同様である。実験結果を図 3.14, 図 3.15, 図 3.16 に示す。また、正確回答の割合を表 3.6 に示す。実験結果や定位精度から分かるように、補正後の定位はほとんど赤直線の上や付近に出ており、データベースの ITD を用い、ITD を補正することによって耳介幅の縮小による定位精度の低下を、前後の誤判定が無く、7 割程度は正確な音像定位が出来る状態まで補正することが可能となったと言える。

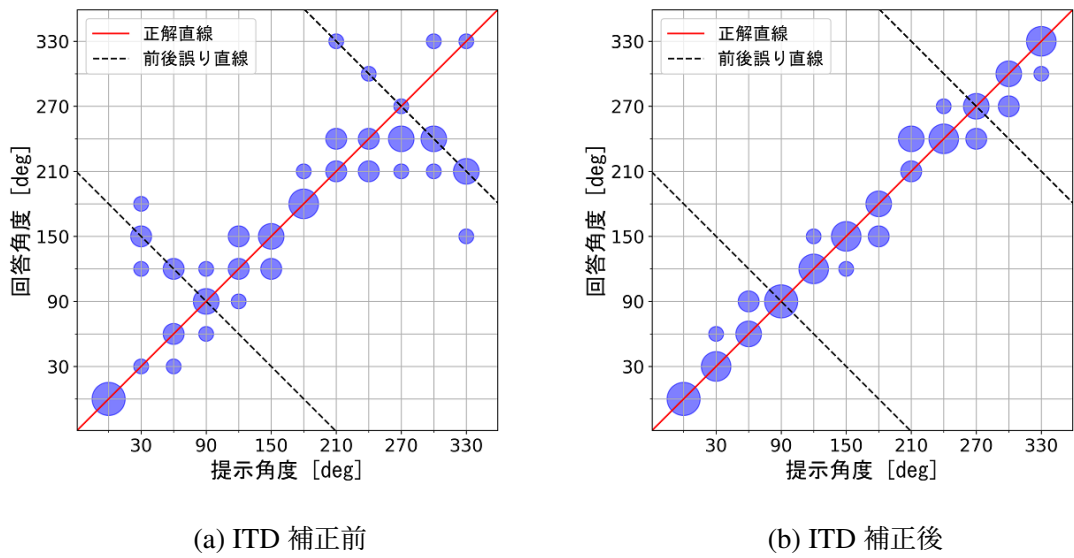
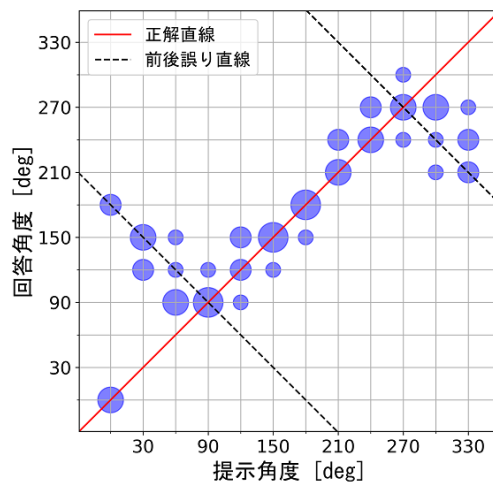
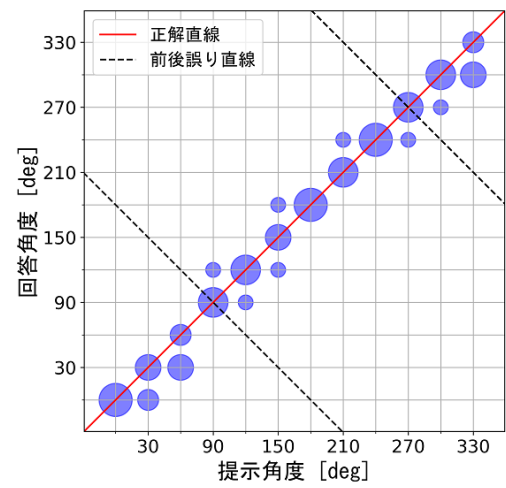


図 3.14: 被験者 A の実験結果

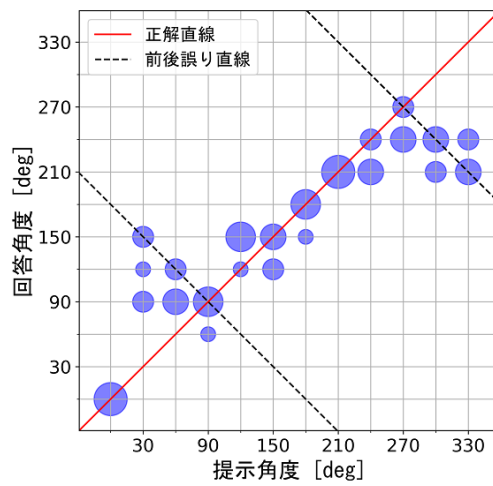


(a) ITD 補正前

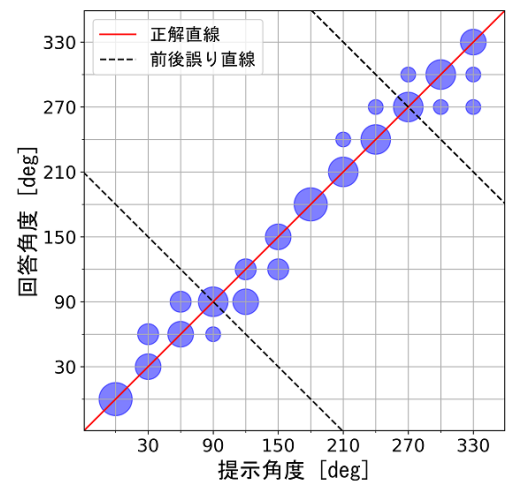


(b) ITD 補正後

図 3.15: 被験者 B の実験結果



(a) ITD 補正前



(b) ITD 補正後

図 3.16: 被験者 C の実験結果

表 3.6: 正確回答割合

被験者	ITD 補正前	ITD 補正後
被験者 A	43.3%	75.0%
被験者 B	41.6%	73.3%
被験者 C	43.3%	73.3%

3.3 まとめ

本章では、提案手法について述べた。3.1 節では、頭部部分の省略として被験者の耳介モデルの作成、それを固定するための耳介幅をダミーヘッドマイクと同様にしたステレオマイクを作成した。また、確認実験としてダミーヘッドマイクに耳介モデルを取り付けたものと作成したステレオマイクに耳介モデルを取り付けたもので被験者に音像定位実験を実施した。結果として、頭部部分の有無によって音像定位精度が変化することは無く、音像定位をさせるうえで頭部部分は省略が可能であることが分かった。

3.2 節では頭部部分の省略によって可能となった両耳介幅の縮小として、耳介モデルは 3.1 節で作成したものをうい、ステレオマイクを出来る限り最小にしたものを作成した。3.1 節と同様、確認実験を実施すると、耳介幅を縮小すると音像定位精度が低下してしまった。音像定位精度低下の原因として、ILD、ITD によるものが考えられ、調査の結果、ITD が原因であった。そこでデータベースの HRIR を用いて ITD の補正を行い定位実験を行ったところ、7 割程度正確な定位をさせることが可能となった。

次章では、提案手法による音像定位精度を実聴の定位精度とを比較し、検討を行う。

第 4 章

提案手法の検証

第 3 章では提案手法として、頭部部分の省略と耳介幅の縮小について述べた。結果として、頭部部分の省略は定位精度に大きな影響は与えず、耳介幅の縮小については ITD の変化を生じさせたが、データベースから算出した ITD によって補正が可能であることが分かった。

本章では、提案手法の検証として、被験者の実聴定位実験の結果との比較を行う。比較を行ったうえで実聴との違いについての調査を行った。

4.1 被験者の実聴時定位実験

4.1.1 実験内容

被験者による実聴の定位試験を実施し、提案手法によって作成したステレオマイクとの音像定位精度の比較を行う。定位実験の内容としては、第 3 章で行ってきた定位実験の収録時と同様に、無響室内において被験者の実聴時の音像定位実験を実施する。

4.1.2 実験条件

無響室での被験者実験の条件は表 4.1 に示す。また、実験環境の概略図を図 4.1 に示す。図のように、被験者の外周に 12 方向 30° 間隔でスピーカーを配置し、ホワイトノイズを聞かせ、音源方向の回答を行ってもらう。スピーカーと被験者の外耳道入りを水平になるよう椅子の高さを調整し、水平面になるようにした。なお、12 方向 5 回ずつの計 60 回の回答を行い、被験者には 12 方向の回答を必ず行ってもらようにした。

表 4.1: 実験条件

被験者	健聴者 3 名
提示音	ホワイトノイズ
角度分解能	30°
音源距離	1.2m
提示回数	60 回 (12 方向 × 5 回)
提示音時間	3sec
スピーカー	MarkAudio PLUVIA7HD-GR-2

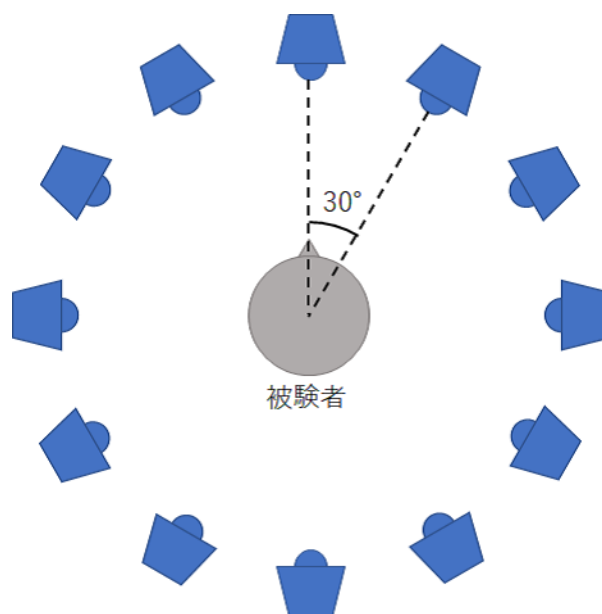


図 4.1: 実験概略図

4.1.3 実験結果

各被験者の実聴定位試験の結果を図 4.2, 図 4.3, 図 4.4 に示す. どの被験者もかなりの高精度で音像定位が出来ており, やはり実聴では大きな定位のズレやブレは見受けられない. 次節では, この実聴での定位試験と提案手法での定位実験の結果, 精度の比較を行う.

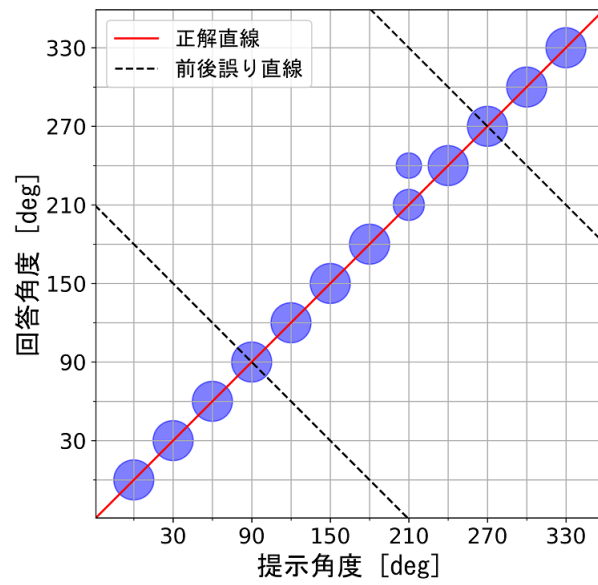


図 4.2: 被験者 A

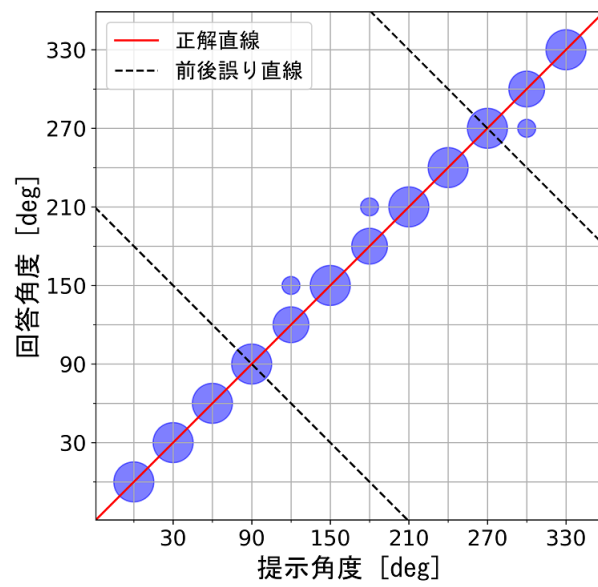


図 4.3: 被験者 B

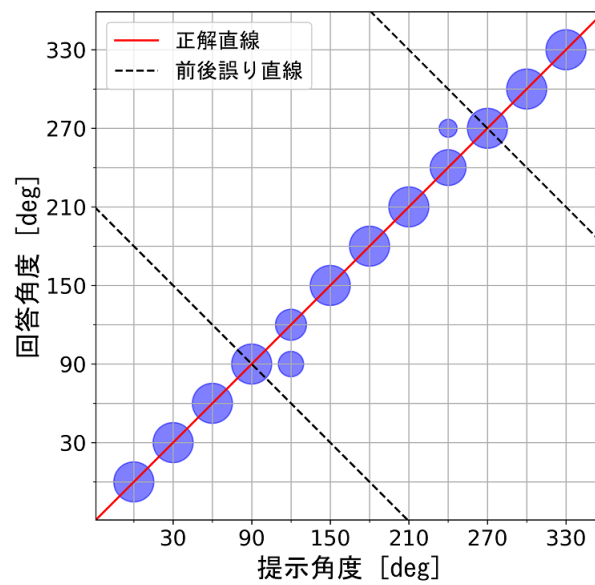
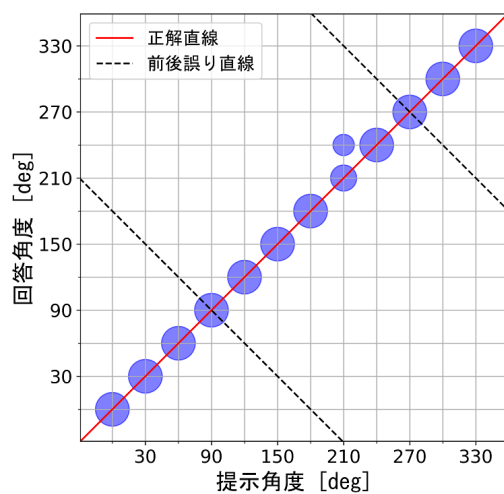


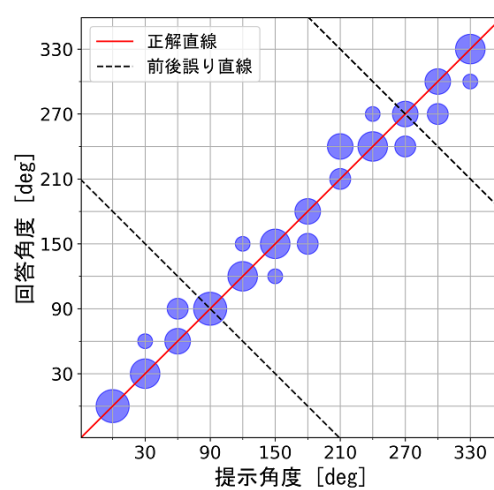
図 4.4: 被験者 C

4.2 提案手法との比較

図 4.5, 図 4.6, 図 4.7 に実聴定位時と提案手法とでの結果の比較を示す. どの被験者も前後誤りは無いが, 実聴定位の結果と比較すると定位にブレがある. また, 表 4.2 に正確定位の比較を示す. これを見ても提案手法は前後誤りこそ無いものの, 実聴の定位と比較すると精度は低くなってしまっていることが分かる. 次節では, この定位精度の低下について原因を調べるために周波数別の定位実験を行う.

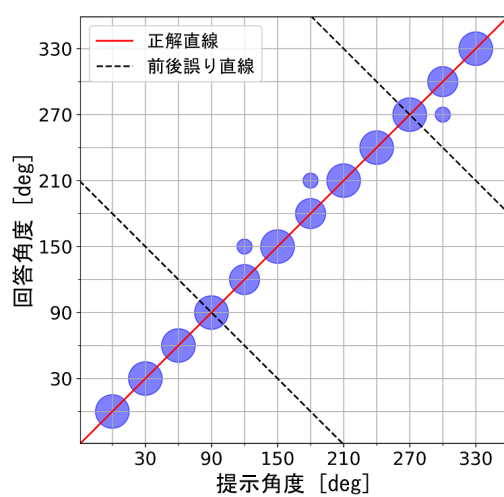


(a) 実聴定位時

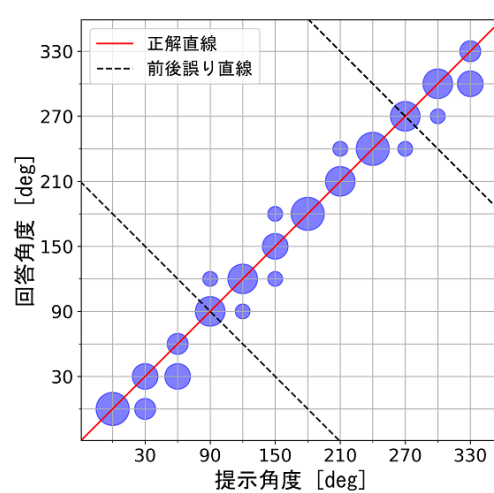


(b) 提案手法

図 4.5: 被験者 A の実験結果

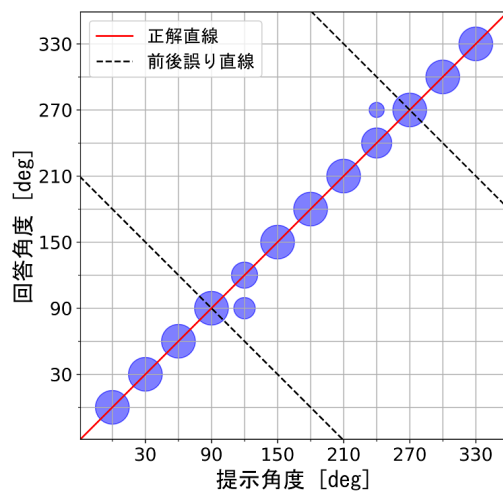


(a) 実聴定位時

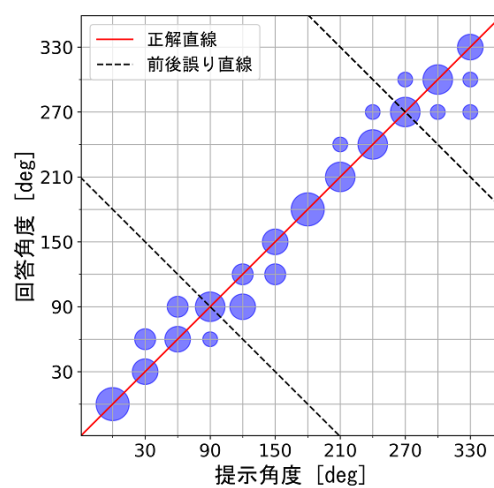


(b) 提案手法

図 4.6: 被験者 B の実験結果



(a) 実聴定位時



(b) 提案手法

図 4.7: 被験者 C の実験結果

表 4.2: 正確回答割合

被験者	実聴定位時	提案手法
被験者 A	96.7%	73.3%
被験者 B	95.0%	75.0%
被験者 C	95.0%	73.3%

4.3 定位精度低下の原因

本節では、実聴定位実験時と比較して提案手法の定位精度が低下してしまった原因を調査する。今回音像定位に使用したホワイトノイズにフィルタをかけて帯域を制限したもので定位実験を行い、どの帯域に原因があるのか突き止め、HRTF の比較を行う。

4.3.1 帯域制限ホワイトノイズによる定位実験

HRTF には周波数帯域によって支配的となっている成分が分かれていることが様々な文献で示されている。今回は、低周波帯域、中周波帯域、高周波帯域の 3 種類の周波数帯域に分け、原因を探る。

今回分けた周波数の詳細を以下に示す。また、フィルタの設計仕様を表 4.3 に示す。

- ・ 低周波帯域
周波数としては 0～2000Hz に LPF をかける。この周波数帯域は両耳間時間差 (ITD: Interaural Time Difference) が支配的な領域となっている。
- ・ 中周波帯域
2000～12000Hz に BPF をかける。この周波数帯域は HRTF のノッチが支配的な領域となっている。
- ・ 高周波帯域
12000～24000Hz に HPF をかける。この周波数には特別支配的な成分は多くないが、2.6 節で述べたように、両耳間レベル差 (ILD: Interaural Level Difference) は高周波帯域になるほど影響があるため、ILD が支配的であると言える。

表 4.3: 各フィルタの設計仕様

フィルタ	カットオフ周波数	使用窓関数	フィルタタップ数
LPF	2000Hz	ハニング窓	255
BPF	2000Hz, 12000Hz	ハニング窓	255
HPF	12000Hz	ハニング窓	255

4.3.1.1 実験条件

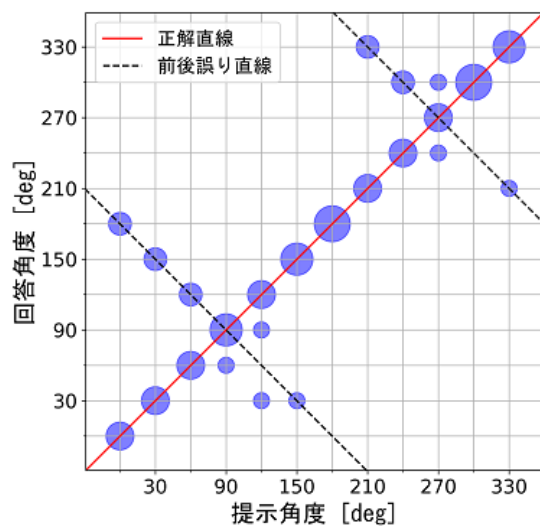
実聴時と提案手法での比較を行うため、被験者には帯域制限を設けたホワイトノイズでの実聴定位実験と作成したステレオマイクを用いての実験を行う。実験条件等は表 3.2、表 4.1 と同様に設定した。

4.3.1.2 実験結果

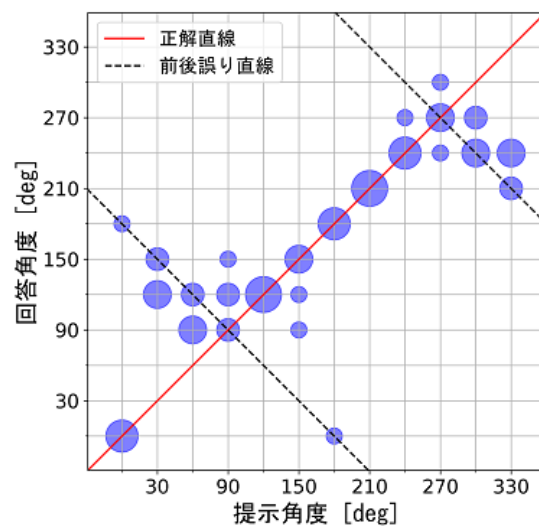
図 4.8, 図 4.9, 図 4.10 に結果を示す。画像枚数が多くなってしまう為、ここでは被験者 1 名分の結果のみを示す。他の被験者の結果は付録を参照していただきたい。低周波帯域の結果を見ると、実聴定位時も定位できている場合もあるが、前後誤りを起こしている場合もあり、定位が出来ているとは言えない。提案手法も同じく定位が出来ておらず、同じような結果と言える。低周波帯域での前後誤りの原因としては、ヒトが前後知覚するために必要である HRTF のノッチの成分が含有されている周波数成分の音が含まれていなかったためであると考えられる。

次に、中周波帯域の結果を見ると、実聴定位時では 4.1.3 節で示したものとほぼ同様な定位精度が出ているのに対して、提案手法の方ではブレが生じてしまっているのが見られる。本節の始めに示したように、中周波帯域は HRTF のノッチが支配的な領域となっており、飯田ら [17] によればこのノッチ部はヒトが前後方向を知覚するのに手がかりとして使用している情報であり、これが多少の変化をしてしまっている可能性が考えられる。これについては次節で述べる。

最後に高周波帯域についてだが、これについても実聴、提案手法ともに前後誤りやバラつきが見られる。また、被験者自身の所感として高周波数は聞こえにくく、定位もしにくかったとの意見があった。これについては等ラウドネス曲線が高周波数において高くなっている、つまりは高周波数の音は音圧を高くしないとヒトには聞こえにくいという特性があり、本実験では条件を揃えるために音圧を揃えていた。これによって聞こえにくくなっていたと考えられる。また、Robert らによると高周波成分は水平面や正中面の音像定位にはあまり寄与しないことを報告している [18] ことから、音像定位精度の低下の原因とは考えにくい。

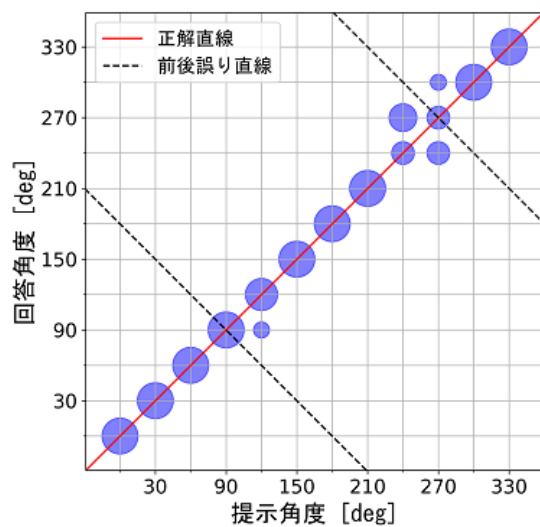


(a) 実聴定位時

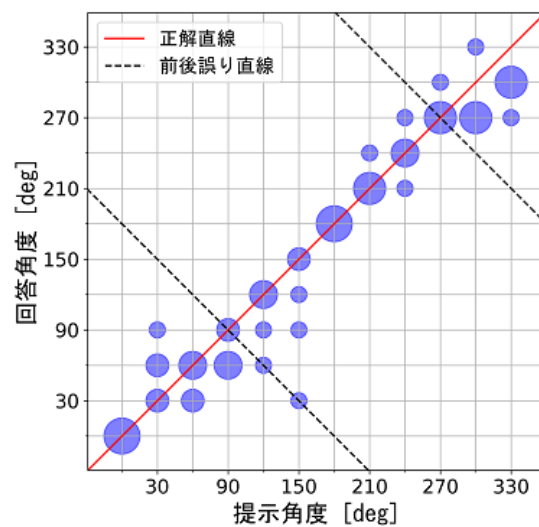


(b) 提案手法

図 4.8: 低周波帯域の定位実験結果



(a) 実聴定位時



(b) 提案手法

図 4.9: 中周波帯域の定位実験結果

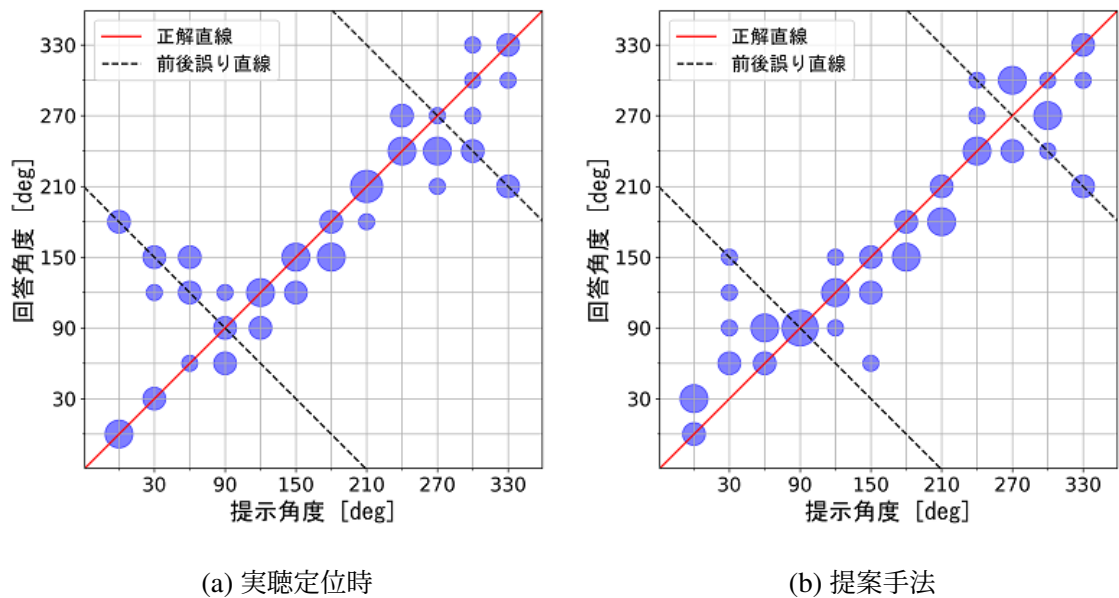


図 4.10: 高周波帯域の定位実験結果

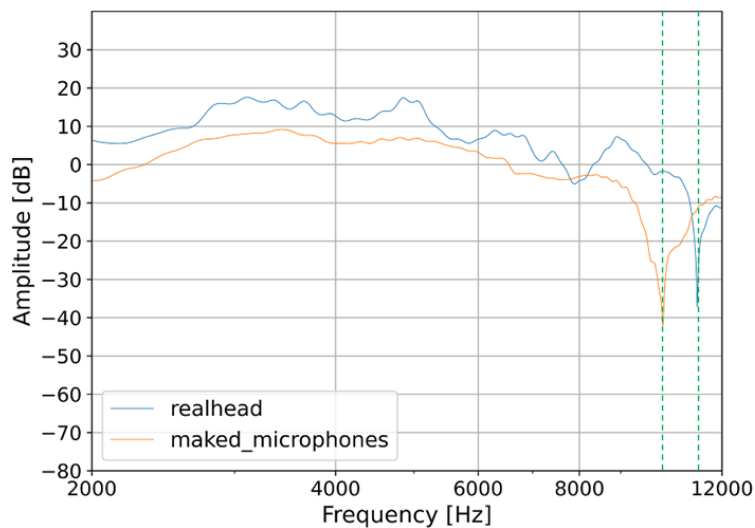
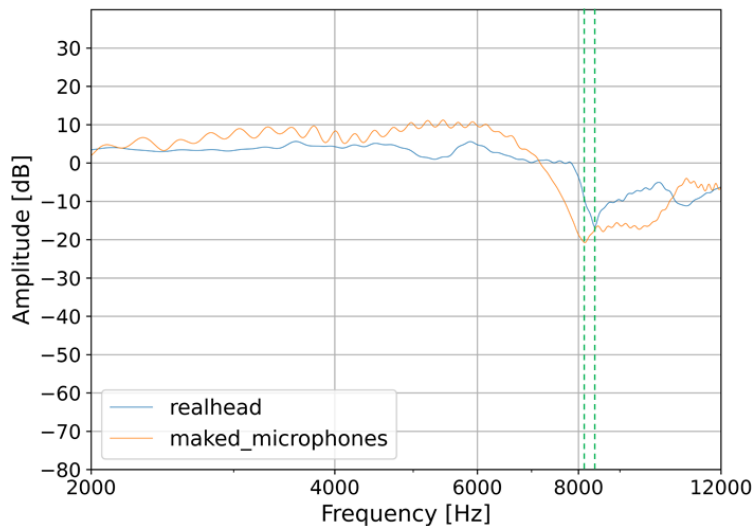
4.3.2 中周波帯域における頭部伝達関数

4.3.1 節では周波数帯域ごとの定位実験の結果を実聴時と提案手法にて比較を行った。結果として中周波帯域の定位精度に違いがあり、中周波帯域にて支配的である HRTF のノッチについて調査する必要があることが分かった。本節では、被験者の実頭部と提案手法で作成したステレオマイクの HRTF 答の特性について調査し、実聴時との定位精度に差がある原因について述べる。

図 4.11, 図 4.12 に実頭部と作成したステレオマイクとの HRTF の比較画像の一例を示す。画像青線は実頭部の HRTF を、橙線は作成したステレオマイクの HRTF を示しており、緑破線はノッチ部分を示している。どちらもノッチの位置にズレが生じており、これによって被験者の定位感にブレが生じてしまい、定位実験の回答方法を 30° 刻みの角度を必ず答える方式にしたため、実験結果にもブレが見られたと考えられる。また、図 4.9 と HRTF の比較を見ると 30° ではノッチ周波数の差が大きいため定位のばらつきも大きくなっているのに対して、 270° ではノッチ周波数の差があまり小さくなく、定位のブレもあまりないように見受けられる。

ノッチ位置の周波数の差について詳しく調査した結果を図 4.13(a), 図 4.13(b) に示す。ここでも被験者 1 名分のみの結果を示す。結果を見ると、反受聴側 (左耳の結果では $30^\circ \sim 150^\circ$, 右耳の結果では $210^\circ \sim 330^\circ$) は周波数の差がかなりあるように見受

けられる。これは、頭部部分の省略によって音の周り込みに差異が生じてしまったためであると考えられる。また、受聴側(左耳の結果では $210^{\circ} \sim 330^{\circ}$ 、右耳の結果では $30^{\circ} \sim 150^{\circ}$)の結果を見ると、 90° 、 270° (正面を 0° としたときに真横に当たる角度)になるにつれてノッチ周波数の差が小さくなっている。真横に当たる部分では外耳道入り口に対して音波が直線的に入ってくると考えられ、ほとんど耳介形状による反射の影響はないので周波数の差が小さく、 60° 、 30° になるにつれて外耳道入り口に入射する音波は耳介形状の影響を受けていると考えられる。以上のことから、受聴側は耳介形状の影響が大きいと考えられる 60° 、 30° になるにつれてノッチ周波数の差が大きくなっていることから、耳介モデルの再現度が不十分であった可能性が考えられる。

図 4.11: 右耳の 30° の HRTF図 4.12: 左耳の 270° の HRTF

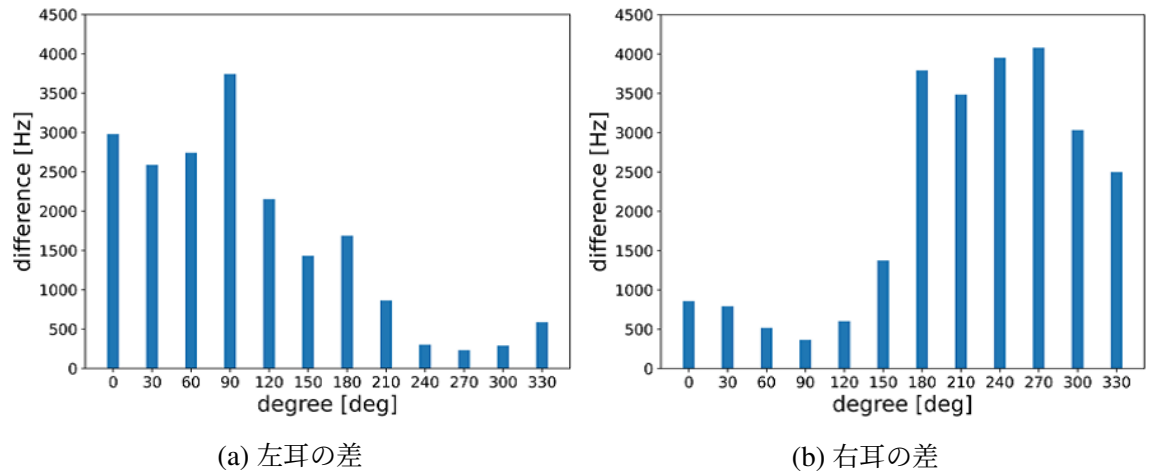


図 4.13: ノッチ位置の周波数の差

4.4 まとめ

本章では、提案手法の有効性の検証として実聴時の定位実験結果との比較を行った。4.1 節で被験者の実聴時の定位試験を行い、4.2 節では提案手法との比較を行った。結果としては、提案手法は実聴の定位時と比較して完全正答割合が低くなっていることが明らかとなった。そこで、4.3 節では定位精度低下の原因の究明のために帯域制限を設けたホワイトノイズによる音像定位実験を行ったところ、中周波帯域に実聴時と提案手法で定位精度に差が見受けられ、それぞれの HRTF を詳しく調査すると、ノッチ位置に差があり、受聴側の 90° 、 270° から離れるにつれて差が大きくなっていったことから、耳介モデルの再現度が不十分ではなかったことが原因として考えられた。しかしながら、提案手法の耳介モデルはハンディタイプの 3D スキャナでスキャンしたものをを用いたモデルであり、MRI などによる精密計測ではない。手軽に計測を行える機器を用いて約 75% の定位精度を出せることを考えると、提案手法は目的をおおよそ達成できていると考えうることが出来る。

第 5 章

おわりに

本論文では、HRIR の測定といった特殊な設備や装置を必要とせず、受聴者の情報収集の負担を最小限にし、小型化を行った取り扱いやすいダミーヘッドシステムの構築を目的とし、従来のダミーヘッドに存在する頭部部分の省略、頭部部分を省略したことによって可能となった両耳介幅の縮小を行い、小型で取り扱いやすいダミーヘッドシステムを提案した。第 1 章では音像定位技術の背景と主な 2 つの手法についてのメリット、デメリットについてと研究の目的について述べた。

第 2 章では、本研究の先行手法として、縮小したダミーヘッドによる音像定位手法と耳介部のみを受聴者本人の耳介モデルを用いて音像定位をさせる手法について述べた。また、本研究の基礎となる知識として頭部伝達関数、両耳間時間差、両耳間レベル差について述べた。

第 3 章では、提案手法である頭部部分の省略、両耳介幅の縮小について、頭部部分の省略に関しては省略の前後で処理を行わずとも同様の結果を得ることが出来た。しかしながら両耳介幅の縮小において、縮小したことによって両耳間時間差が変化してしまい、定位精度が低下してしまった。両耳間時間差の補正方法としてデータベースにある情報のみを用いて補正を行ったところ、定位精度の補正につながった。

第 4 章では被験者の実聴時との定位精度の比較を行った。比較の結果、提案手法の方が定位精度が低いことが分かった。原因の解明のために、周波数帯域制限を設けたホワイトノイズによる音像定位実験を行ったところ、中周波帯域に実聴時と提案手法での差が見られた。中周波帯域は HRTF のノッチが支配的な領域であり、この部分に問題があると考えた。HRTF を比較した結果ノッチの位置に差があり、受聴側の真横の角度に当たる 90° 、 270° から離れるにつれてノッチ位置の差が大きくなっていった。このことから、耳介モデルの再現度が不十分であったことが考えられる。しかしながら、今回は MRI などの精密計測ではなく、3D スキャナによる簡易的な計測であり、それによって約 75% の定位精度を出すことが出来ており、エンターテイメント分野で

の応用といった形での実用は十分であると考えられる。

今後の課題として、まず正中面に対する検討が挙げられる。今回の手法は水平面のみに対する手法の検討であり、正中面に関しての検討は行っていない。より臨場感のある音を提供するうえでは音の左右だけでなく上下の情報も重要であり、検討する必要性はあると考えられる。

また、今回の定位精度では医療や福祉の分野への応用は難しいと考える。定位精度向上の手掛かりとして、作成する耳介モデルの精度を向上させることが考えられる。しかし、精密なモデルの作成には MRI 等による精密な計測が必要であると考えられ、受聴者への計測時の負担は本手法と比較すると大きくなってしまうと考えられる。

謝辞

本研究を行うにあたって、適切にご指導ご鞭撻頂いた、杉田泰則准教授に深く感謝致します。また、本論文の執筆にあたり、副査をご担当いただき、多くの助言を賜りました本学の電気系岩橋政宏教授ならびに圓道知博准教授、原川良介助教に感謝致します。そして、苦楽を共にした信号処理応用研究室の皆様感謝いたします。協力して頂いた皆様へ心から感謝の気持ちとお礼を申し上げたく、謝辞にかえさせていただきます。

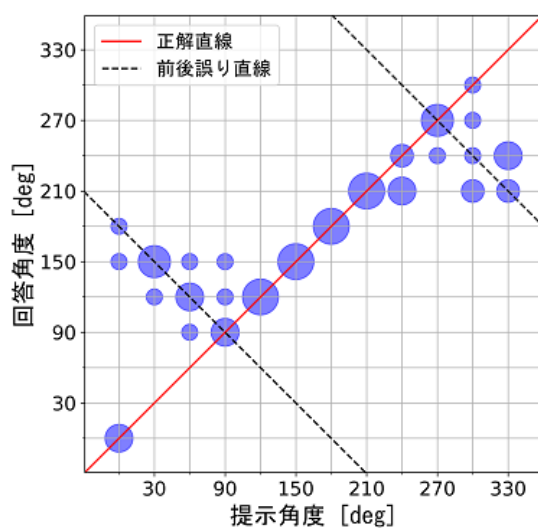
最後に、勉学に励む機会を与えて下さり、様々な面で支えていただいた両親に深く感謝いたします。

令和4年2月

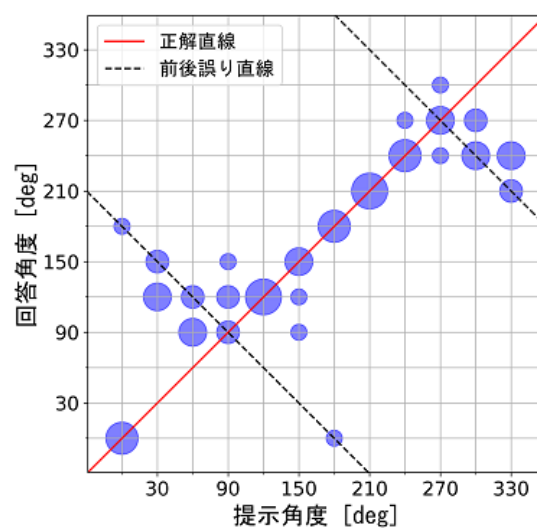
付録 A

他被験者の実験結果

A.1 他被験者の帯域制限音像定位実験結果

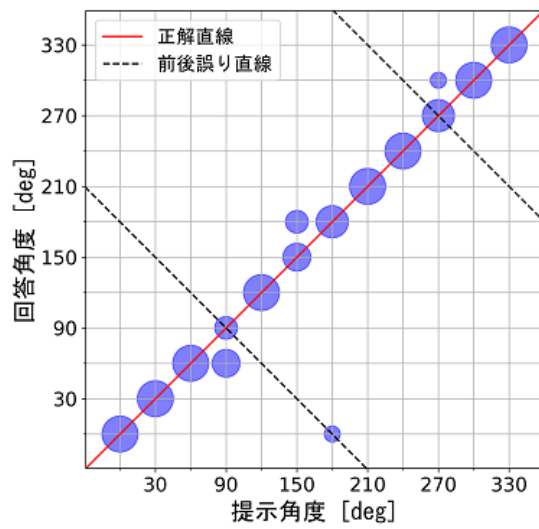


(a) 実聴定位時

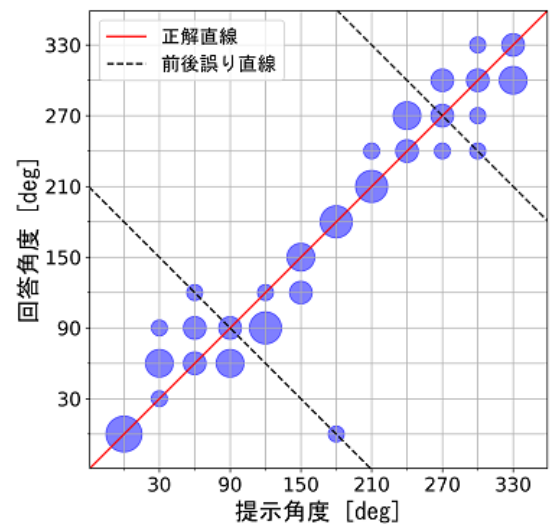


(b) 提案手法

図 1.1: 被験者 B の低周波帯域の定位実験結果

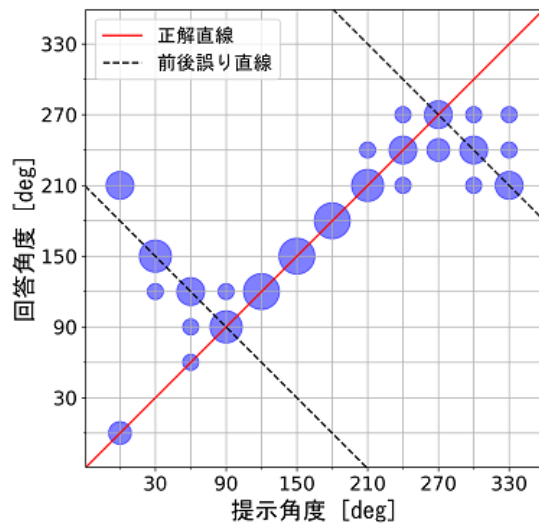


(a) 実聴定位時

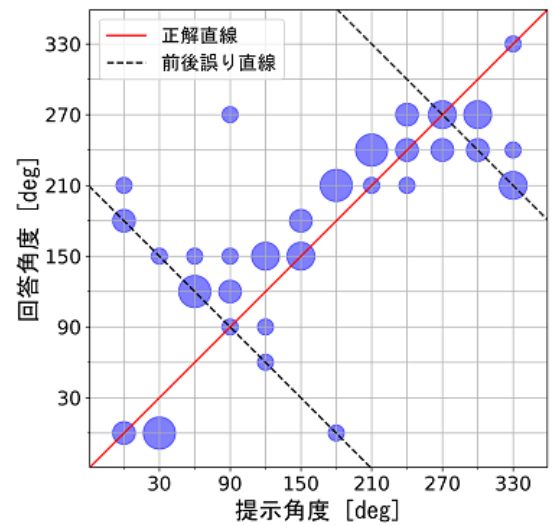


(b) 提案手法

図 1.2: 被験者 B の中周波帯域の定位実験結果

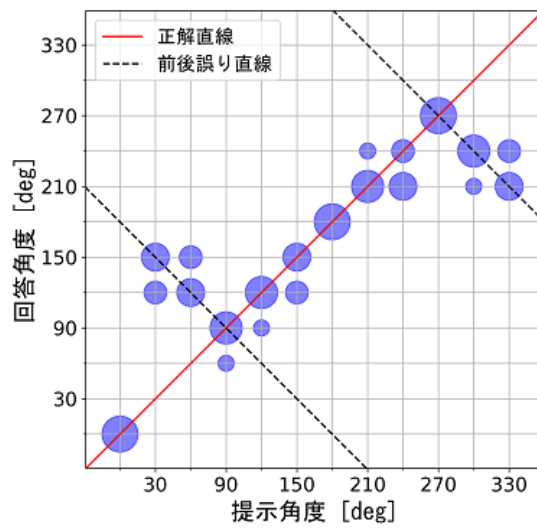


(a) 実聴定位時

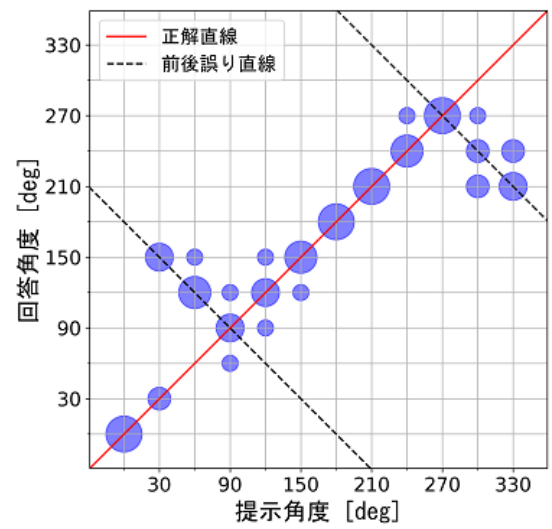


(b) 提案手法

図 1.3: 被験者 B の高周波帯域の定位実験結果

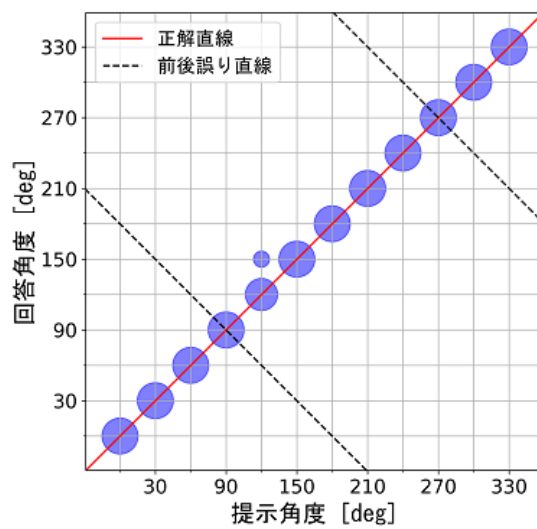


(a) 実聴定位時

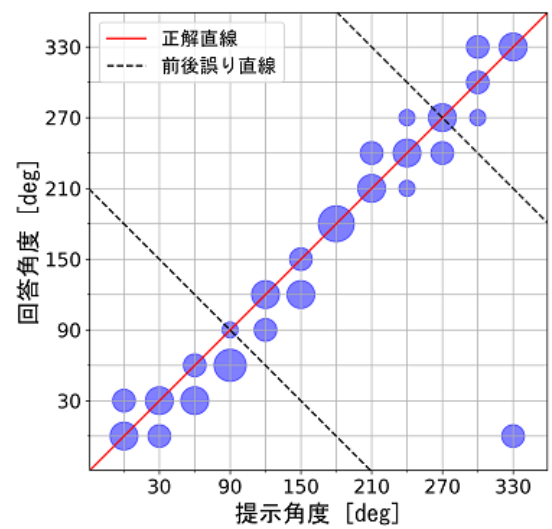


(b) 提案手法

図 1.4: 被験者 C の低周波帯域の定位実験結果

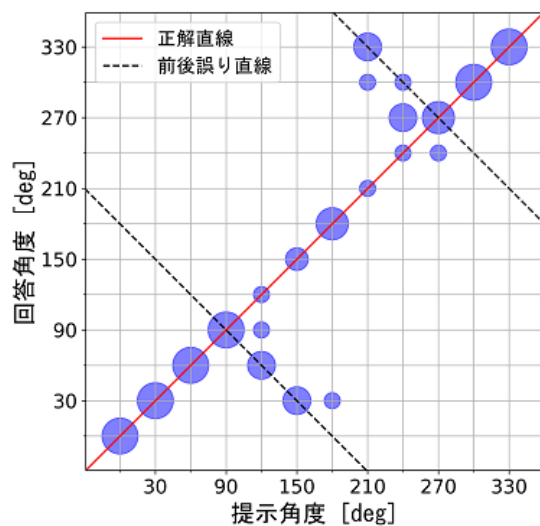


(a) 実聴定位時

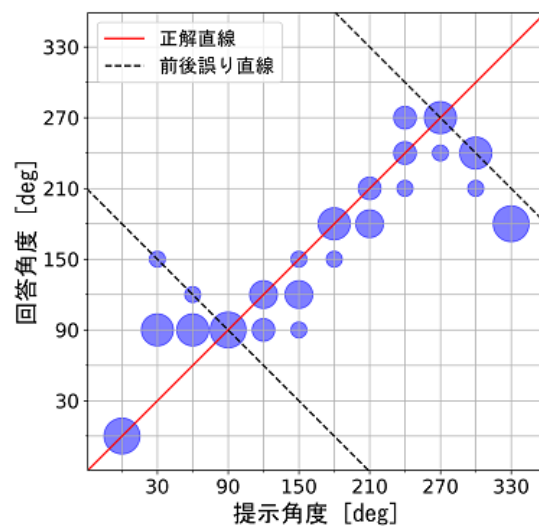


(b) 提案手法

図 1.5: 被験者 C の中周波帯域の定位実験結果



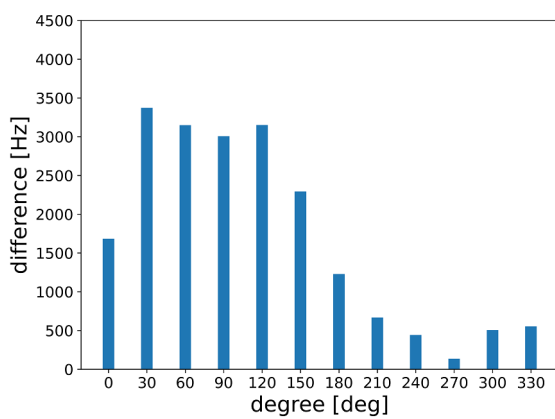
(a) 実聴定位時



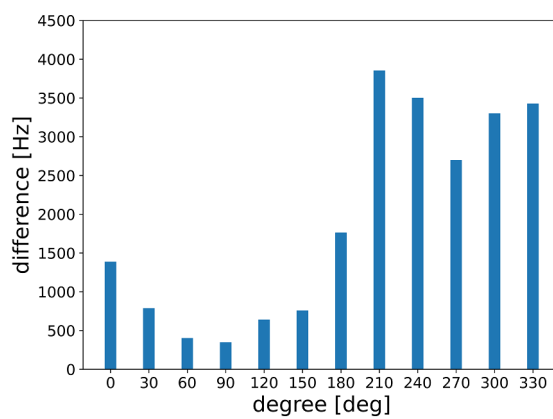
(b) 提案手法

図 1.6: 被験者 C の高周波帯域の定位実験結果

A.2 他被験者のノッチ位置の周波数差



(a) 左耳の差



(b) 右耳の差

図 1.7: 被験者 B のノッチ位置の周波数の差

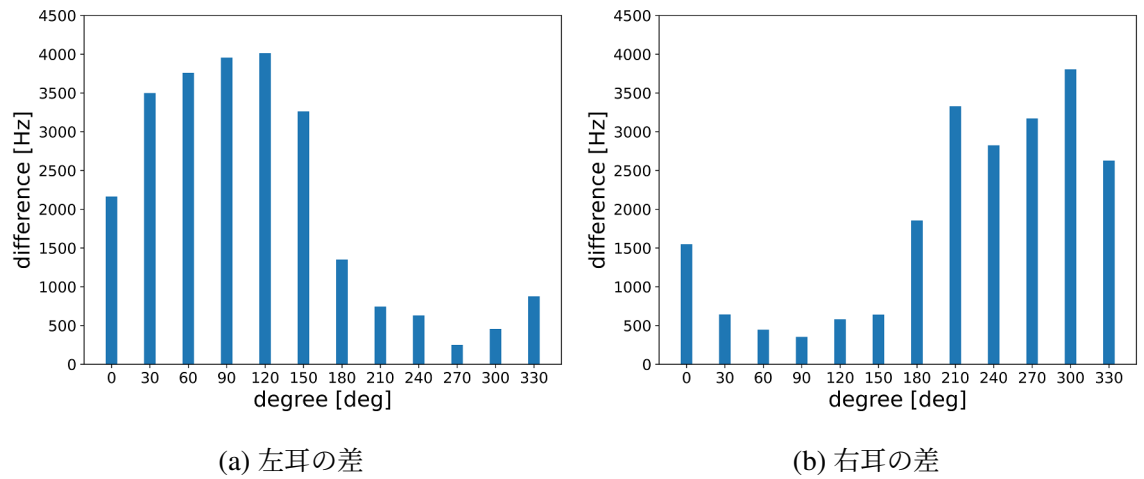


図 1.8: 被験者 C のノッチ位置の周波数の差

参考文献

- [1] 飯田一博. 頭部伝達関数の個人性. 日本音響学会誌, Vol. 73, No. 3, pp. 181–188, 2017.
- [2] 林阿希子, 伊藤亜紗, 渡邊淳司. スポーツ・ソーシャル・ビュー：競技を身体的に翻訳し視覚障がい者と共有する生成的スポーツ観戦手法. 日本バーチャルリアリティ学会論文誌, Vol. 25, No. 3, pp. 216–227, 2020.
- [3] 関喜一. 音響 vr 技術を用いた視覚障害者の聴覚訓練. *BME:bio medical engineering*, Vol. 12, No. 8, pp. 38–43, 1998.
- [4] 岩谷幸雄. 頭部伝達関数による音像定位. 日本音響学会誌, Vol. 73, No. 3, pp. 173–180, 2017.
- [5] J Blauert. Spacial hearing: The psychophysics of human sound localization, 1997.
- [6] K.A.J.Riederere. Repeatability analysis of head-related transfer function measurements. *journal of the audio Engineering Society*, Vol. 105, No. 4846, 1998.
- [7] E.M.Wenzel, M.Anuda, D.J.Kistler, and F.L.Wightman. Localization using nonindividualized head - related transfer functions. *J.Acoust. Soc. Am.*, Vol. 120, pp. 111–123, 1993.
- [8] 平原達也, 相良洋行, 遠藤渉. 縮小ダミーヘッドによる音像定位. 日本音響学会誌講演論文集, pp. 433–434, 2006.
- [9] 中村俊一, 小見茂光, 子安勝. 録音用人工頭の製作における耳介の個人的形状の重要性について. 日本音響学会誌, Vol. 32, No. 10, pp. 618–620, 1976.
- [10] 戸嶋巖樹, 青木茂樹. 音響テレプレゼンスロボットにおける頭部形状簡略化の音響的・知覚的影響. 人工知能学会, Vol. 31, No. 8, pp. 788–796, 2013.
- [11] E.A.G.Shaw and R.Teranishi. Sound pressure generated in an external-ear replica and real human ears by a nearby point source. *J.Acoust.Soc.Am.*, Vol. 44, pp. 240–249, 1968.
- [12] B.F.G.Katz and M.Noisternig. A comparative study of interaural time delay estimation methods. *J.Acoust.Soc.Am.*, Vol. 135, pp. 3530–3540, 2014.
- [13] 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-両耳間時

- 間差の推定-. 日本バーチャルリアリティ学会論文誌, Vol. 22, No. 3, pp. 404–412, 2017.
- [14] 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. 電子情報通信学会技術研究報告, Vol. 116, No. 54, pp. 1–12, 2016.
- [15] R.G. Klumpp and H.R. Eady. Some measurements of interaural time difference thresholds. *The Journal of the Acoustical Society of America*, Vol. 28, pp. 859–860, 1956.
- [16] Kanji Watanabe, Yukio Iwara, Yoiti Suzuki, Shouichi Takane, and Sojun Sato. Dataset of head-related transfer functions measured with a circular loudspeaker array. *Acoust. Sci. and Tech*, Vol. 35, No. 3, 2014.
- [17] Kazuhiro Iida, Motokuni Itoh, Atsue Itagaki, and Masayuki Morimoto. Median plane localization using parametric model of the head-related transfer function based on spectral cues. *Applied Acoustics*, Vol. 63, No. 8, pp. 835–850, 2007.
- [18] B Robert and R Simon. The impact of signal bandwidth on auditory localization: implications for the design of three-dimensional audio displays. *Human Factors*, Vol. 39, No. 2, pp. 287–295, 1997.