

長岡技術科学大学大学院
工学研究科修士論文

題 目

骨伝導ヘッドホンによる
音像定位精度向上に関する研究

指導教員 準教授 杉田 泰則

著者 電気電子情報工学専攻
11106284 多田 紘大

提出期日 平成29年2月13日

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 本論文の構成	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 [10]	3
2.1.1 座標系	3
2.1.2 頭部伝達関数	4
2.1.3 バイノーラル音源	5
2.2 歪み成分耳音響放射 (Distortion Product OtoAcoustic Emission, DPOAE) [8]	6
2.3 気道音・骨導音の振幅特性差の推定手法 [9]	8
2.4 まとめ	12
第 3 章 提案手法	13
3.1 伝達特性差	13
3.2 提案手法 1 (「音の再生位置の違い」の解決)	17
3.3 提案手法 2 (「クロストーク」の除去)	20
3.4 まとめ	23
第 4 章 提案手法の検証	24
4.1 バイノーラル音源の測定条件	24
4.2 気道音・骨導音の振幅特性差の推定	26
4.3 被験者評価試験	30
4.4 まとめ	34

第 5 章 おわりに	35
謝辞	36
参考文献	38

第1章 はじめに

1.1 研究背景

音像定位技術とは、ヘッドホンなどの受聴音に対して方向感などの位置情報を付与する技術である。この技術は映画やゲームなどのエンターテイメント分野をはじめ、視覚障害者向けの歩行支援機器などの医療福祉分野など様々な分野への応用が期待されている [1][2]。

一般に気道ヘッドホンを用いた音像定位において、頭部伝達関数 (Head-Related Transfer Function, HRTF) を組み合わせることで、高い音像定位精度が実現可能である。HRTF とは音源から外耳道入り口までの伝達関数であり、頭や胴体、耳介などに起因したものである。

近年、骨伝導ヘッドホンを用いた音像定位技術に関する研究が行われている [3]-[6]。骨伝導ヘッドホンは振動子によって頭蓋骨を加振し、その振動が骨を介して蝸牛に到達することで音を受聴するものである。故に、気道ヘッドホンのように耳を塞がないため、周囲の環境音を受聴することが可能である。また、骨伝導による音の受聴は気道音による受聴よりも疲労感が少ないため、長時間の使用に適しているという利点もある [7]。

しかし、気道ヘッドホンと HRTF を組み合わせることで高い音像定位精度を実現できるのに対し、骨伝導ヘッドホンと HRTF を組み合わせた場合には定位精度が低いのが現状である。定位精度低下の大きな要因として以下の 2 つが考えられる。1 つ目は「音の再生位置の違い」である。気道音が外耳道から中耳を経由し蝸牛へ到達するのに対して、骨導音は振動子から骨を経由して蝸牛へ到達する。HRTF は音源から外耳道入り口までの伝達関数であるため、HRTF を畳み込んだ信号は外耳道入り口での音の再生を前提としている。そのため、骨伝導ヘッドホンでは音の再生位置が異なってしまい、音像定位精度の低下を招く要因となる。2 つ目は「クロストーク」の影響である。気道ヘッドホンの場合には左（右）の再生音が左（右）の蝸牛にのみ到達する。しかし、骨伝導ヘッドホンは左右の振動子によって頭蓋骨を加振しているため、右側で加振した振動が左側の蝸牛に、左側で加振した振動が右側の蝸牛に到達する。これをクロストークと呼び、音像定位精度の低下を招く要因となる。骨伝導ヘッドホンと HRTF の組み合わせによる音像定位精度向上のためには、上記の問題を解決する必要がある。

1.2 研究目的

本論文では、骨伝導ヘッドホンによる音像定位精度向上を目的とし、「音の再生位置の違い」と「クロストーク」の問題を解決する2つの前処理手法を提案する。提案法では、外耳道入り口から蝸牛までの伝達特性に対する、振動子から蝸牛までの伝達特性の差を利用する。伝達特性の差の推定には、生体反応である歪み成分耳音響放射（Distortion Product OtoAcoustic Emission, DPOAE）を用いる[8]。具体的にはDPOAE発生用の2つの刺激音の片方に音圧の変化を与え、それを気道で与えた場合と骨道で与えた場合のDPOAEを比較することで推定を行う[9]。気道ヘッドホン、前処理無しの骨伝導ヘッドホン、前処理を適用した骨伝導ヘッドホンのそれぞれについて被験者評価試験を実施し、音像定位精度への影響を検証する。

1.3 本論文の構成

本論文の構成は以下の通りである。第2章では本論文で用いる基礎的知識について述べる。また、提案手法は伝達特性差を用いるが、この伝達特性差の推定の基礎となる、外耳道入り口から蝸牛までの振幅特性に対する、振動子から蝸牛までの振幅特性の差の推定手法について述べる。第3章では第2章で述べた手法に基づく、「音の再生位置の違い」と「クロストーク」の問題を解決する2つの提案手法について述べる。第4章では被験者評価試験の結果を示す。被験者評価試験では水平面30刻みでバイノーラル録音した白色雑音を用いて、提案法が音像定位精度に与える影響を検証する。第5章ではこれまでに述べた内容、結果をまとめ、結論を述べる。

第 2 章 本研究の基礎となる知識・手法

この章では、本研究で用いる基礎的な知識及び、提案手法に用いる伝達特性差の推定の基礎となる、振幅特性差の推定手法について説明する。まずははじめに、頭部伝達関数を用いた音像定位技術について述べる。次に、生体反応である歪み成分耳音響放射について述べる。最後に振幅特性差の推定手法について述べる。

2.1 音像定位技術 [10]

ヘッドホンなどから受聴する音に方向感などの位置情報を与え、受聴者に音像を知覚させる技術を音像定位技術と呼ぶ。気道ヘッドホンを用いて、高い精度の音像定位技術を実現するためには、頭部伝達関数を用いる手法が一般的である。以下の 2.1.2 節、2.1.3 節において頭部伝達関数を用いた音像定位技術について述べる。

2.1.1 座標系

まずははじめに、本論文において音源の方向を表すための座標系について述べる。本論文では、図 2.1 に示す、頭を中心とした球座標系を用いる。この座標系の原点は両外耳道入り口を結ぶ線分の中点である。眼窩点（眼窓の入っている穴の下縁）と耳珠（外耳道入り口の出っ張り）を含む平面を水平面と呼ぶ。また、両外耳道入り口を通り、水平面に垂直な面を横断面と呼ぶ。正中面は、水平面と横断面に対して垂直な面である。角度 θ は方位角、角度 ϕ は仰角と呼ばれている。

本論文では、音源が水平面上に存在する場合に対する検証を行っているため、仰角 $\phi = 0^\circ$ であり、これ以降 ϕ は省略する。

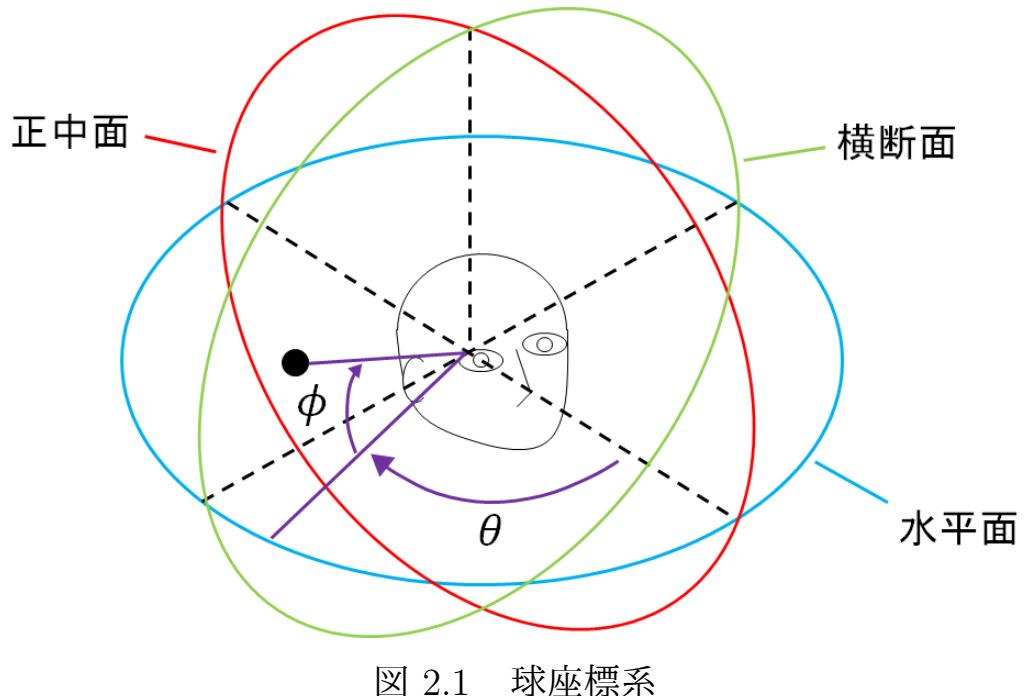


図 2.1 球座標系

2.1.2 頭部伝達関数

図 2.2 に示すように, ある音源から放射された音がヒトの耳に到達するとき, 受聴者の頭や胴体, 耳介の形状などによる反射・回折の影響を受ける. これらの影響による伝達特性を表したもののが頭部伝達関数 (Head-Related Transfer Function, HRTF) であり, 次式で定義される.

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

ここで, $H_{L|R}(\omega, \theta)$ 左右の HRTF であり, ω は周波数, θ は音源の方位角である. $G_{L|R}(\omega, \theta)$ は音源から左右それぞれの外耳道入り口までの伝達関数である. $F(\omega, \theta)$ は受聴者がいない状態での, 音源から頭部中心の位置までの伝達関数である.

HRTF は受聴者に対する音源の角度によって特性が異なり, その特性の違いからヒトは音源の方向を知覚している. また, HRTF には個人差があり, 他者の HRTF を用いて音像定位を行う場合, 前後の不一致などが頻繁に発生し, 定位精度が低下してしまう [11].

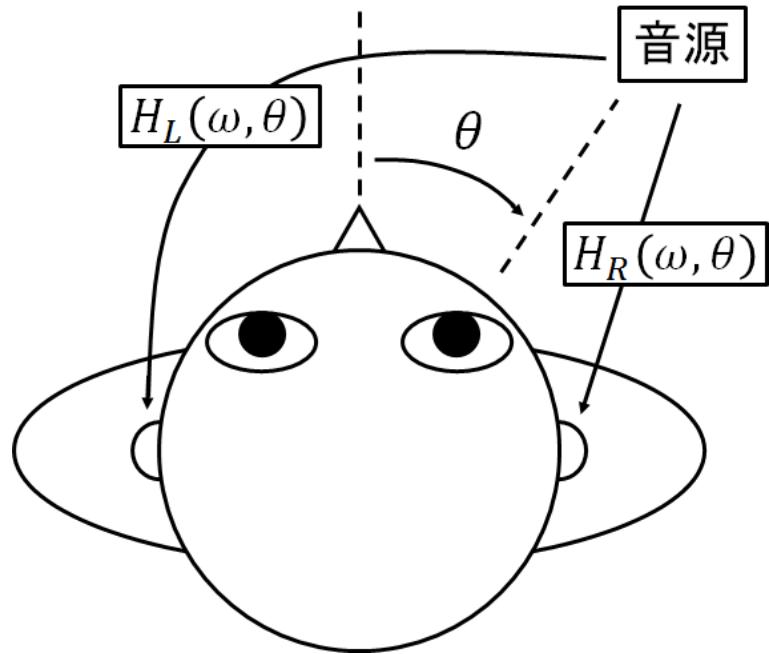


図 2.2 音の受聴

2.1.3 バイノーラル音源

2.1.2 節で述べたように、音がヒトの耳に到達するとき、HRTF の影響を受ける。この HRTF の影響を受けた、耳に到達する信号をバイノーラル音源と呼び、次式で表すことができる。

$$X_{L|R}(\omega, \theta) = H_{L|R}(\omega, \theta)S(\omega) \quad (2.2)$$

ここで、 $X_{L|R}(\omega, \theta)$ は左右の耳におけるバイノーラル音源であり、 $S(\omega)$ は音源である。このバイノーラル音源は、音源から放射され外耳道入り口に到達する信号であり、ヒトはこれを手がかりに音源の位置情報を知覚している。すなわち、このバイノーラル音源を気道ヘッドホンなどの音を再生する機器と組み合わせることで、実際に音源から放射された際の受聴音を再現することが可能であり、高い音像定位精度を実現することができる。

2.2 歪み成分耳音響放射 (Distortion Product OtoAcoustic Emission, DPOAE) [8]

歪み成分耳音響放射 (Distortion Product OtoAcoustic Emission, DPOAE) とは、ヒトの生体反応のひとつであり、耳音響放射 (OtoAcoustic Emission, OAE) の一種である。図 2.3 にヒトの耳の構造を示す。通常ヒトが音を受聴する際、信号は外耳から中耳を経由し、内耳にある蝸牛へと到達し電気信号へと変換され脳が音を知覚する。それに対して、OAE とは蝸牛において生成された信号が、中耳を経由し外耳方向へと伝播・放射される信号の総称である。内耳に存在する蝸牛の内部はリンパ液で満たされており、外耳から入力された音、すなわち空気の振動は蝸牛でリンパ液の振動に変換される。このリンパ液の振動を拾い、電気信号へと変換するのが蝸牛内部に存在する外有毛細胞である。OAE はこの外有毛細胞の働きによって生成される信号である。

この OAE には複数の種類が確認されており、それぞれ発生要因ごとに分類がされている。最も大きな分類として、自発耳音響放射 (Spontaneous OtoAcoustic Emission, SOAE) と誘発耳音響放射 (Evoked OtoAcoustic Emission, EOAE) に分けられる。SOAE とは、自発的に発生する OAE の総称であり、刺激音などによって発生する種類の OAE とは違い、常時内耳にて発生し、外耳方向へと放射されている。EOAE は SOAE とは違い、刺激音によって生成される OAE の総称である。EOAE はさらに短音刺激誘発耳音響放射 (Transient Evoked OtoAcoustic Emission, TEOAE) と、本研究に用いる DPOAE に分類ができる。図 2.4 に簡単な OAE の分類図を示す。

DPOAE とは、十分な振幅を持った周波数の異なる 2 つの刺激音を耳に入力することで、蝸牛の非線形性によって生成される信号である。内耳で生成され外耳方向に放射される信号は、2 つの周波数の和音や差音、結合音を含んでいる。DPOAE は蝸牛内の外有毛細胞の働きを評価するのに適しており、一般的に行われる主観的な聴覚検査を行うことが困難な、乳幼児などの聴覚検査に用いられている。

ここで 2 つの刺激音の周波数をそれぞれ f_1, f_2 ($f_1 < f_2$) とする。放射される DPOAE の大きさは 2 つの刺激音の周波数に関係しており、周波数比が $f_2/f_1 = 1.22$ 付近のときに最も大きく放射されることが報告されている [12]。また、放射される DPOAE のうち、 $2f_1 - f_2$ 周波数成分が、他の非線形線分に対して最も大きく観測可能であることが確認されている [13]。さらに、刺激音の音圧が変化することによって DPOAE の音圧も変化するが、その変化は非線形的であることが分かっている [14]。

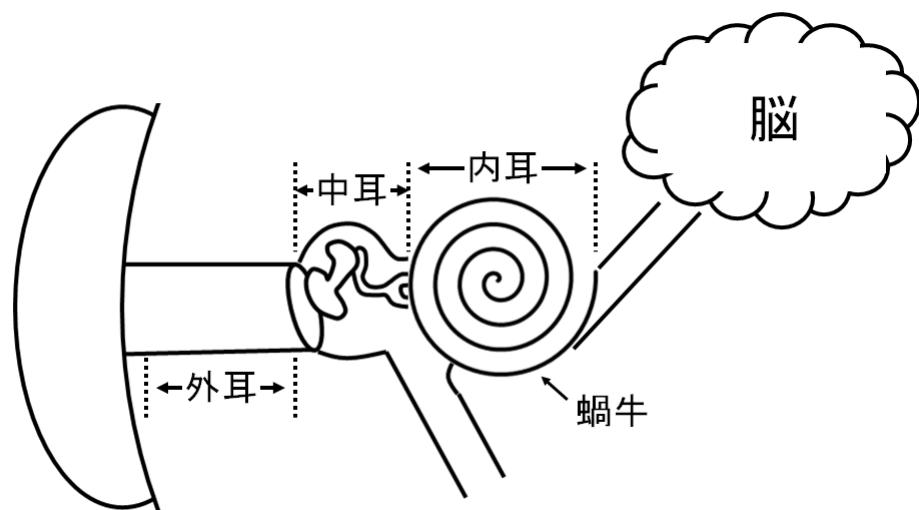


図 2.3 耳の構造

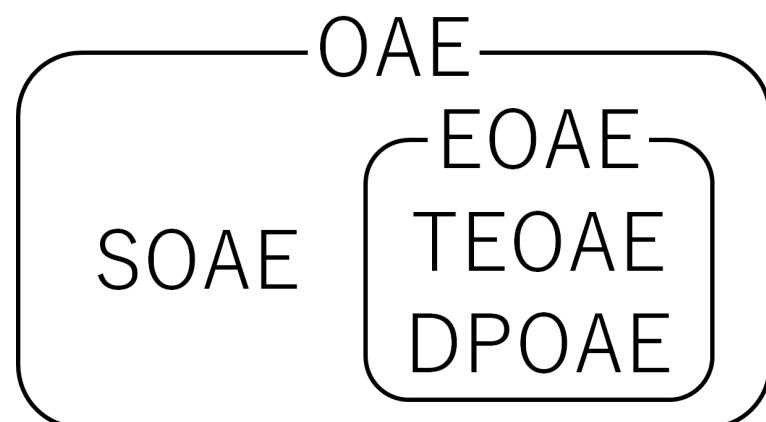


図 2.4 OAE の分類

2.3 気道音・骨導音の振幅特性差の推定手法 [9]

DPOAE を用いて、外耳道入り口から蝸牛までの振幅特性に対する、骨伝導振動子から蝸牛までの振幅特性の差を推定する手法が提案されている。2.2節において述べた DPOAE は 2 つの刺激音によって生成される信号であるが、これは骨導音で与えた刺激音を用いても生成される。この手法では、図 2.5 に示すように 2 つの刺激音を気道音で与えた場合の DPOAE と、図 2.6 に示すように気道音・骨導音それぞれで与えた場合の DPOAE を比較することで、振幅特性の差を推定する。

ここで、DPOAE 発生用の 2 つの刺激音をそれぞれ s_1, s_2 とする。また、それぞれの周波数を f_1, f_2 ($f_1 < f_2$)、音圧を L_1, L_2 とする。

まずははじめに、 s_1, s_2 を共に気道音で与えた場合の DPOAE を測定する。このとき、 L_1 は一定値に固定し、 L_2 を複数の値に変化させた場合の測定を行う。刺激音の音圧が変化することで DPOAE の音圧も変化し、図 2.7 のようなグラフを得ることができる。図 2.7 の縦軸は DPOAE の音圧 ($2f_1 - f_2$ 周波数成分) であり、横軸は s_2 の音圧 L_2 となっている。例では、測定条件は $f_1 = 4918\text{Hz}$, $f_2 = 6000\text{Hz}$, $L_1 = 75\text{dB}$ であり、 L_2 は 45dB から 75dB まで 5dB 間隔で変化させている。

次に、 s_1 を気道音、 s_2 を骨導音で与え、上記と同様の測定を行う。図 2.8 に測定結果を示す。 L_2 を 55dB から 85dB の範囲で変化させてある。その他の条件に関しては、気道音の場合と同様である。

次に、2 つの測定によって得られた結果から、振幅特性差の推定を行う。図 2.7 と図 2.8 を合わせたグラフを、図 2.9 に示す。図 2.9 より、どちらのグラフも山なりの概形を持っており、片方のグラフを横軸方向にシフトすることで、おおよそ重なることがわかる。ここで s_2 を骨導音で与えたときのグラフを 0.1dB 間隔で横軸方向にシフトさせ、2 つのグラフの二乗誤差が最小になるようなシフト量を算出する。2 つのグラフは L_2 を 5dB 間隔で測定したものであるため、 0.1dB 間隔でシフトさせるために、線形補間により間の値を算出する。シフト後のグラフを図 2.10 に示す。この例では、 s_2 を骨導音で与えたときのグラフを -6.3dB シフトした際に、自乗誤差が最小になった。つまり、周波数 6000Hz (f_2 周波数)において、骨導音は気道音に対して 6.3dB 減衰して蝸牛に到達するといえる。

このように各周波数において、上記測定を行うことで気道音に対する、骨導音の振幅特性の差を推定する事ができる。

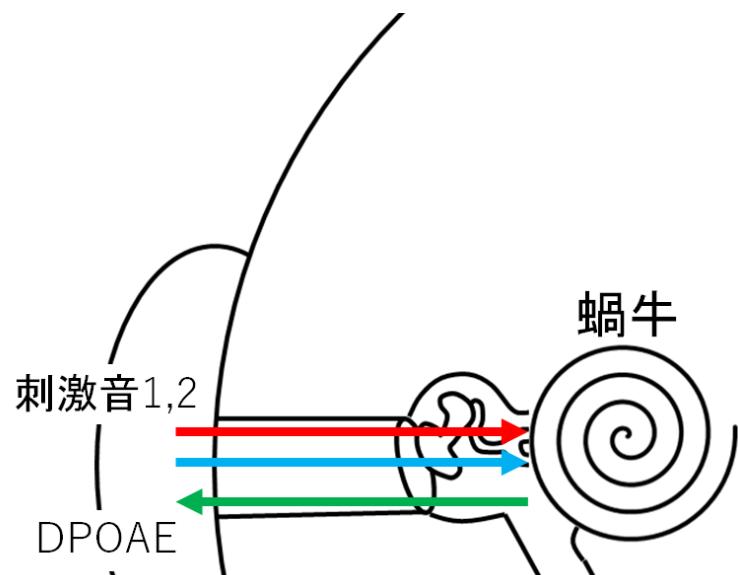


図 2.5 気道音による DPOAE

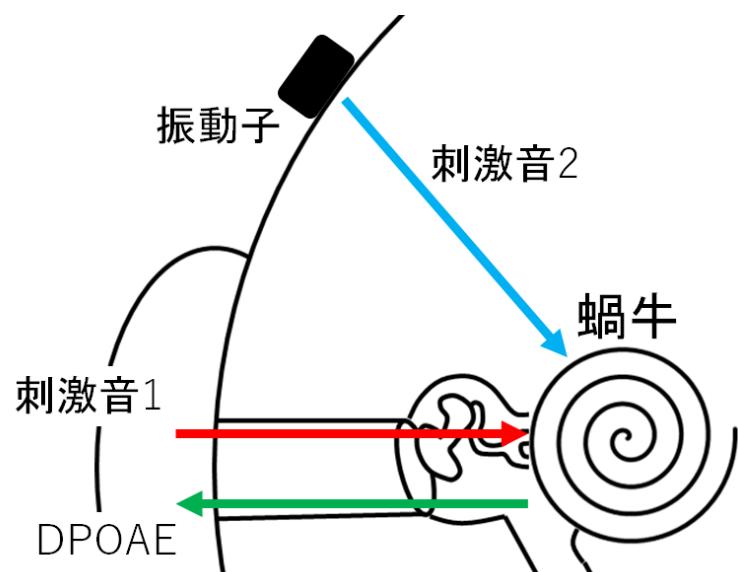


図 2.6 骨導音による DPOAE

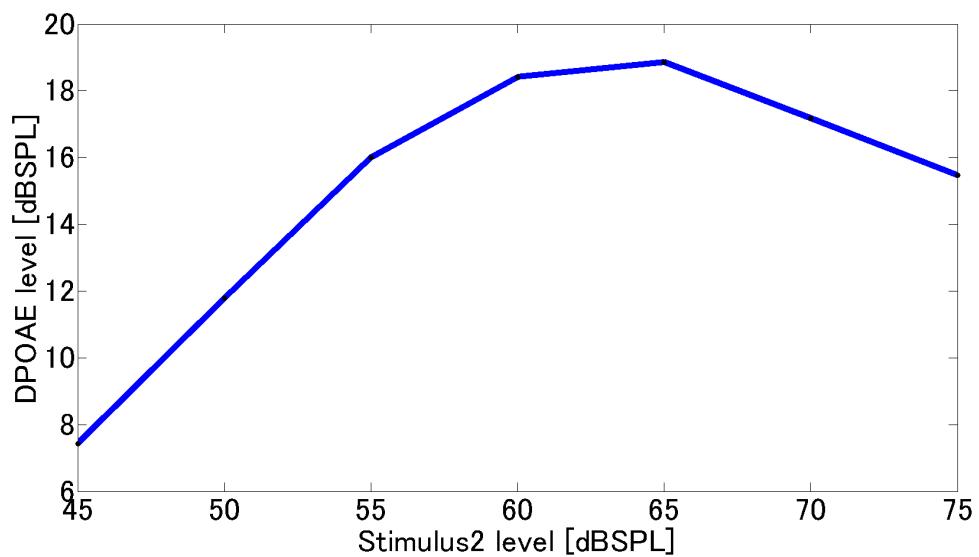


図 2.7 気道音変化による DPOAE

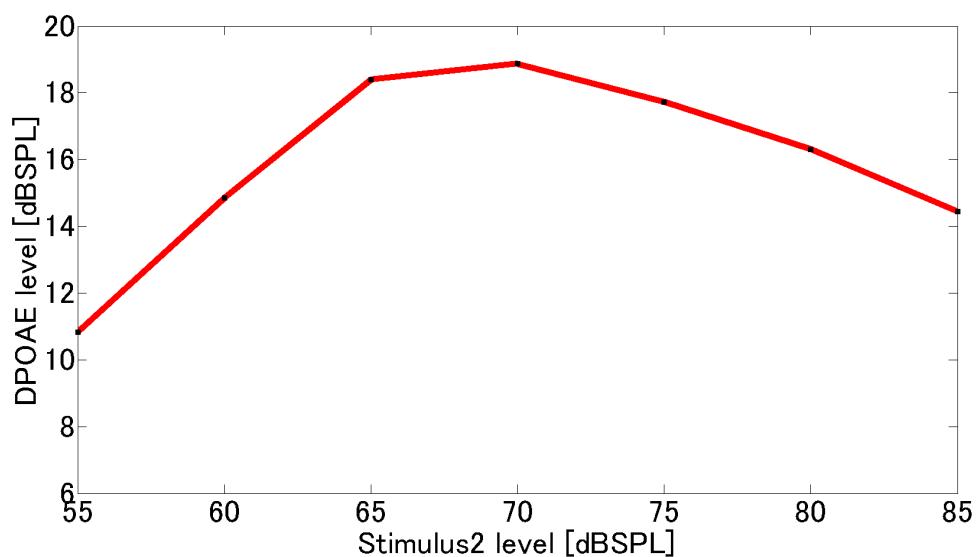


図 2.8 骨導音変化による DPOAE

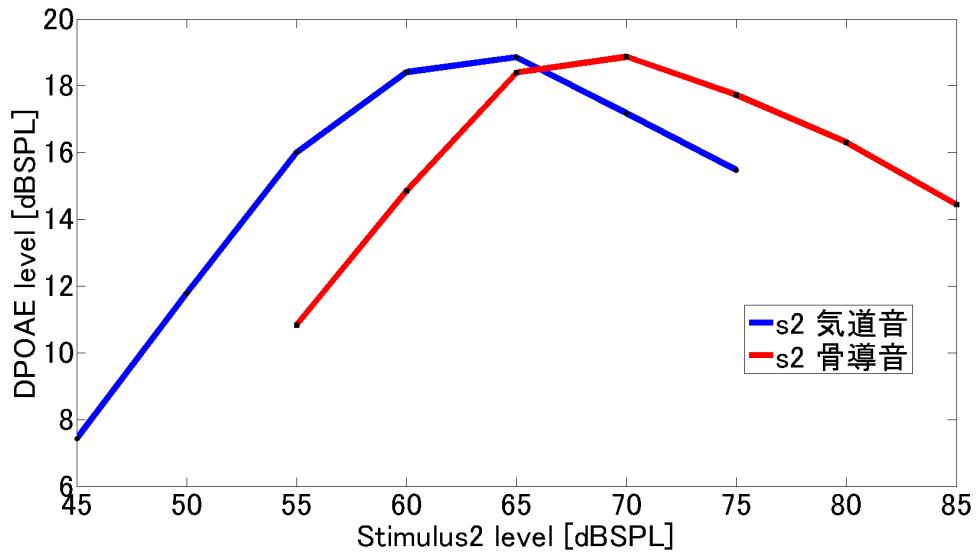


図 2.9 気道音・骨導音変化による DPOAE

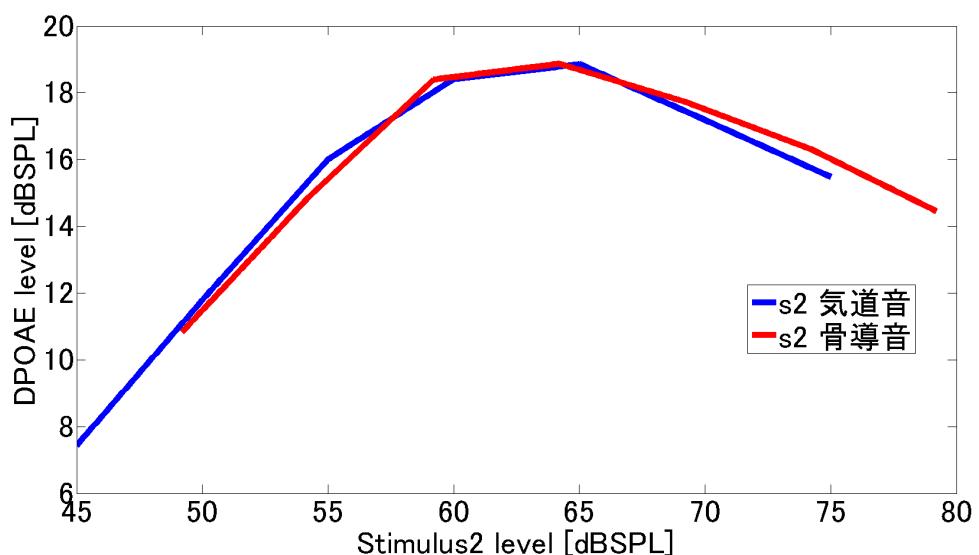


図 2.10 気道音・骨導音変化による DPOAE (シフト後)

2.4 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。ヒトが音を受聴する際に、ある音源から放射された信号が外耳道入り口に到達するまでの伝達関数がHRTFである。また、左右の外耳道入り口に到達する信号をバイノーラル音源とよぶ。これはHRTFの影響を受けた信号であるため、気道ヘッドホンなどの外耳道入り口で音を再生する機器と組み合わせることで、実際の受聴音を再現することが可能である。OAEは生体反応のひとつであり、発生要因ごとに分類がされている。その中の一種であるDPOAEを用いた、気道音・骨導音の振幅特性差の推定手法が提案されている。次章では、提案手法である、骨伝導ヘッドホンによる音像定位精度向上のための、伝達特性差を用いた前処理フィルタについて述べる。

第 3 章 提案手法

この章では、外耳道入り口から蝸牛までの伝達特性に対する、骨伝導振動子から蝸牛までの伝達特性の差を用いた、提案手法を示す。まず、3.1節において、提案手法に用いる伝達特性差について述べる。次に、3.2節において、「音の再生位置の違い」を解決する前処理手法を示す。最後の3.3節において、「クロストーク」の問題を解決する前処理手法を示す。

3.1 伝達特性差

本研究の提案手法では、気道音と骨導音の伝達特性差を用いる。ここで、音の伝達経路を図3.1に示す。また、対応する記号を表3.1に示す。

表 3.1 伝達経路に対応する記号の説明

記号	説明
$H_{ACR}(z)$	右の外耳道入り口から右の蝸牛までの気道伝達関数
$H_{ACL}(z)$	左の外耳道入り口から左の蝸牛までの気道伝達関数
$H_{BCRR}(z)$	右の振動子から右の蝸牛までの骨導伝達関数
$H_{BCLL}(z)$	左の振動子から左の蝸牛までの骨導伝達関数
$H_{BCRL}(z)$	右の振動子から左の蝸牛までの骨導伝達関数
$H_{BCLR}(z)$	左の振動子から右の蝸牛までの骨導伝達関数

提案手法では気道音に対する骨導音の伝達特性差、すなわち $H_{BCRR}(z)/H_{ACR}(z)$, $H_{BCLL}(z)/H_{ACL}(z)$, $H_{BCLR}(z)/H_{ACR}(z)$, $H_{BCRL}(z)/H_{ACL}(z)$ を用いる。そこで、2.3節の推定により得られるものは振幅特性差であるため、本研究では位相特性差に関しては線形位相を仮定し、上記伝達特性差を実現する FIR フィルタを最小自乗法を用いて作成する。

測定によって、周波数軸上で L 点の離散的な振幅特性差が推定されたとする。ここで、この振幅特性を実現する N 次の FIR フィルタを作成することを考える。 $h(n)$ をフィルタ係数として、式(3.1) に FIR フィルタの式を示す。

$$H(\omega) = \sum_{n=0}^N h(n) e^{-j\omega n} \quad (3.1)$$

また, L 点の実現したい周波数特性を $D(\omega_l)$ と置く. 振幅特性 (推定により得られる) を $|D(\omega_l)|$, 位相特性 (線形位相という仮定により得られる) を $\theta(\omega_l)$ とすると以下のように表せる.

$$D(\omega_l) = |D(\omega_l)|e^{j\theta(\omega_l)} \quad (3.2)$$

ただし, 位相特性は線形位相を仮定しているため $\theta(\omega_l) = -(N/2)\omega_l$ である. この, 所望の特性を実現する FIR フィルタを作成するため, 式 (3.1) に示した FIR フィルタとの自乗誤差を計算する. 自乗誤差 E を, 離散化した周波数点における近似関数と所望特性によつて表すと以下のようにになる.

$$E = \sum_{l=0}^{L-1} |H(\omega_l) - D(\omega_l)|^2 \quad (3.3)$$

ここで $\omega_l = 2\pi l/L$ である. 最小自乗法では, この誤差 E が最小になるようなフィルタ係数 $h(n)$ を作成する手法である. 例として, 図 3.2 に示す離散的な振幅特性差について, 最小自乗法によりこの特性を実現する 256 次の FIR フィルタを作成した. 所望の特性と, 作成した FIR フィルタの振幅特性を重ねたグラフを図 3.3 に示す.

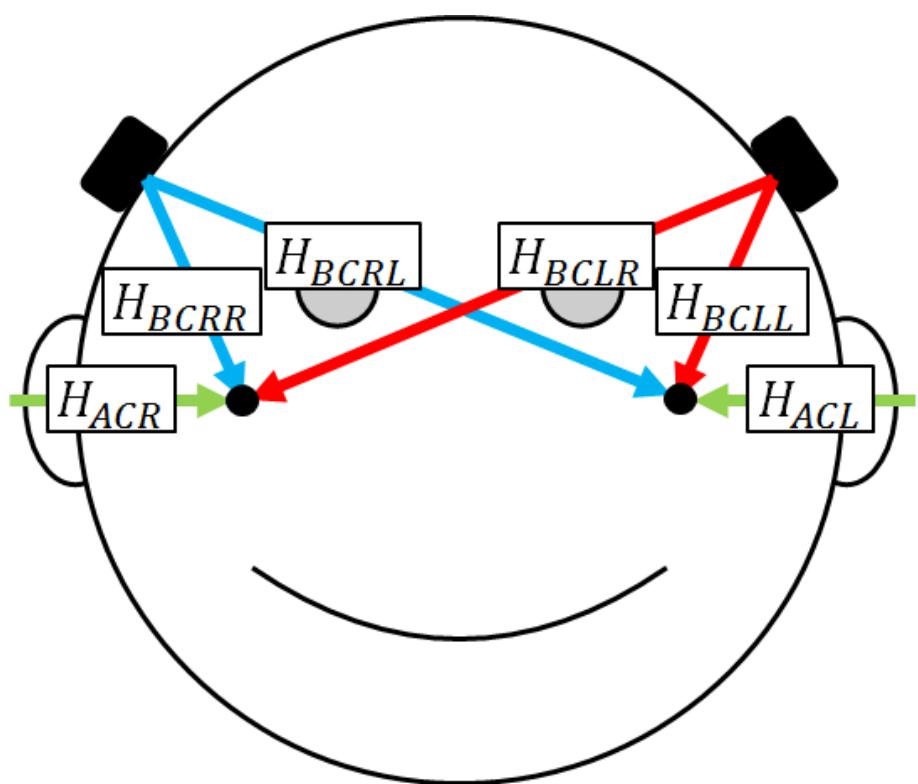


図 3.1 音の伝達経路

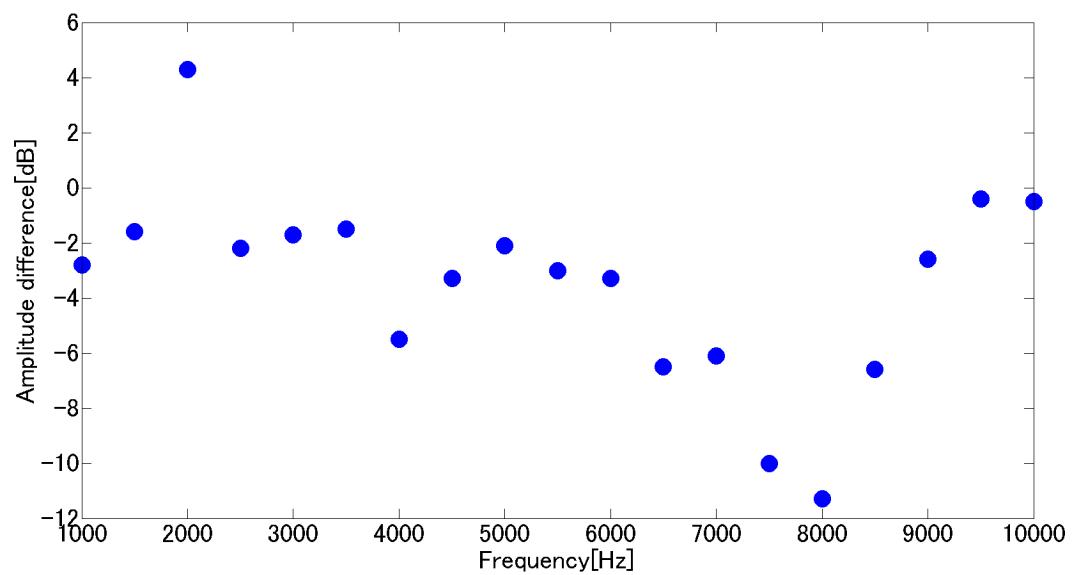


図 3.2 離散的な振幅特性差

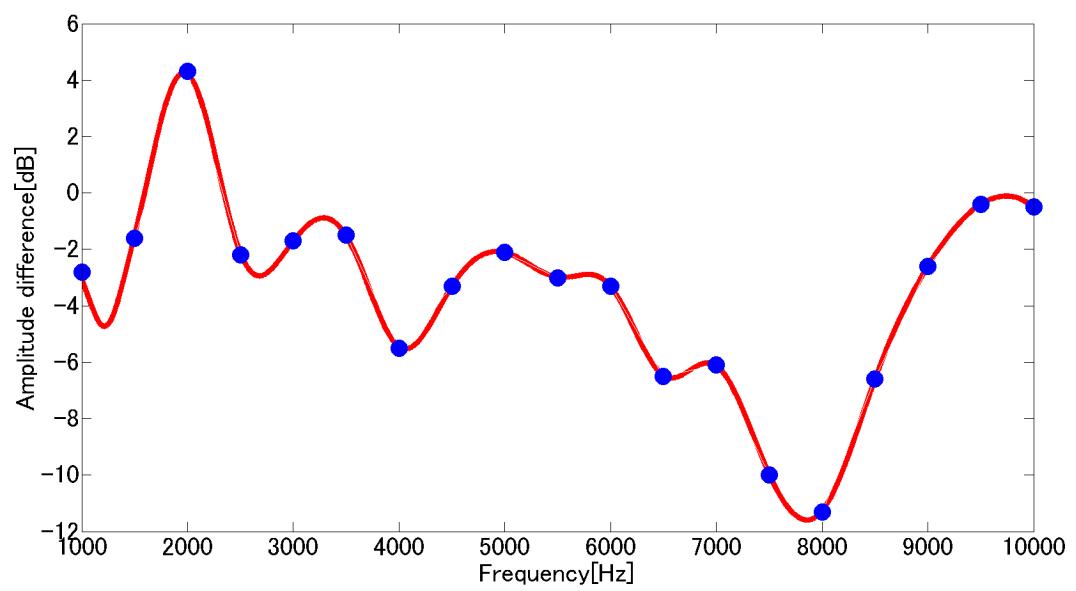


図 3.3 所望の特性を実現するフィルタ特性

3.2 提案手法1（「音の再生位置の違い」の解決）

1つ目の提案手法は、骨伝導ヘッドホンとHRTFの組み合わせによる音像定位技術における、「音の再生位置の違い」を解決する手法である。この手法では、図3.4に示すように2つの前処理フィルタ $G_R(z)$, $G_L(z)$ を用いる。ここで、外耳道入り口で再生される信号を $x_R(t)$, $x_L(t)$ とし、外耳・中耳を経由し蝸牛に到達する信号を $z_{ACR}(t)$, $z_{ACL}(t)$ とする。また、それぞれ z 変換したものを、 $X_R(z)$, $X_L(z)$, $Z_{ACR}(z)$, $Z_{ACL}(z)$ とすると、以下のようになる。

$$Z_{ACR}(z) = X_R(z)H_{ACR}(z) \quad (3.4)$$

$$Z_{ACL}(z) = X_L(z)H_{ACL}(z) \quad (3.5)$$

また、左右の振動子において再生される信号を $y_R(t)$, $y_L(t)$ 、骨を経由して蝸牛に到達する信号を $z_{BCR}(t)$, $z_{BCL}(t)$ とする。それぞれ z 変換したものを、 $Y_R(z)$, $Y_L(z)$, $Z_{BCR}(z)$, $Z_{BCL}(z)$ とすると、以下のようになる。

$$Z_{BCR}(z) = Y_R(z)H_{BCRR}(z) \quad (3.6)$$

$$Z_{BCL}(z) = Y_L(z)H_{BCLL}(z) \quad (3.7)$$

この提案手法において、 $Y_R(z)$, $Y_L(z)$ は、 $X_R(z)$, $X_L(z)$ が前処理フィルタ $G_R(z)$, $G_L(z)$ を通ったものであるため、式(3.6)及び、式(3.7)は以下のようになる。

$$Z_{BCR}(z) = X_R(z)G_R(z)H_{BCRR}(z) \quad (3.8)$$

$$Z_{BCL}(z) = X_L(z)G_L(z)H_{BCLL}(z) \quad (3.9)$$

今、Rchの信号について考える。ここで、式(3.4)の気道音 $Z_{ACR}(z)$ と、式(3.8)の骨導音 $Z_{BCR}(z)$ を等価にする前処理フィルタ $G_R(z)$ を考える。このとき $G_R(z)$ は以下のように求めることができる。

$$\begin{aligned}
Z_{ACR}(z) &= Z_{BCR}(z) \\
X_R(z)H_{ACR}(z) &= X_R(z)G_R(z)H_{BCRR}(z) \\
G_R(z) &= \left(\frac{H_{BCRR}(z)}{H_{ACR}(z)} \right)^{-1}
\end{aligned} \tag{3.10}$$

同様に Lch について、式 (3.5) の気道音 $Z_{ACL}(z)$ と、式 (3.9) の骨導音 $Z_{BCL}(z)$ を等価にする前処理フィルタ $G_L(z)$ は以下のようになる。

$$G_L(z) = \left(\frac{H_{BCLL}(z)}{H_{ACL}(z)} \right)^{-1} \tag{3.11}$$

よって、前処理フィルタを用いて気道音と骨導音を等価にするには、 $(H_{BCRR}(z)/H_{ACR}(z))^{-1}$ 及び、 $(H_{BCLL}(z)/H_{ACL}(z))^{-1}$ が必要である。

ここで、 $(H_{BCRR}(z)/H_{ACR}(z))$ と $(H_{BCLL}(z)/H_{ACL}(z))$ は、前節において述べた通り、2.3 節の振幅特性差推定手法と位相特性差が線形位相という仮定を用いて推定可能である。よってこの推定手法により得られた伝達特性差を実現するフィルタの逆フィルタを設計することで、提案手法の前処理フィルタ $G_R(z)$, $G_L(z)$ を実現できる。

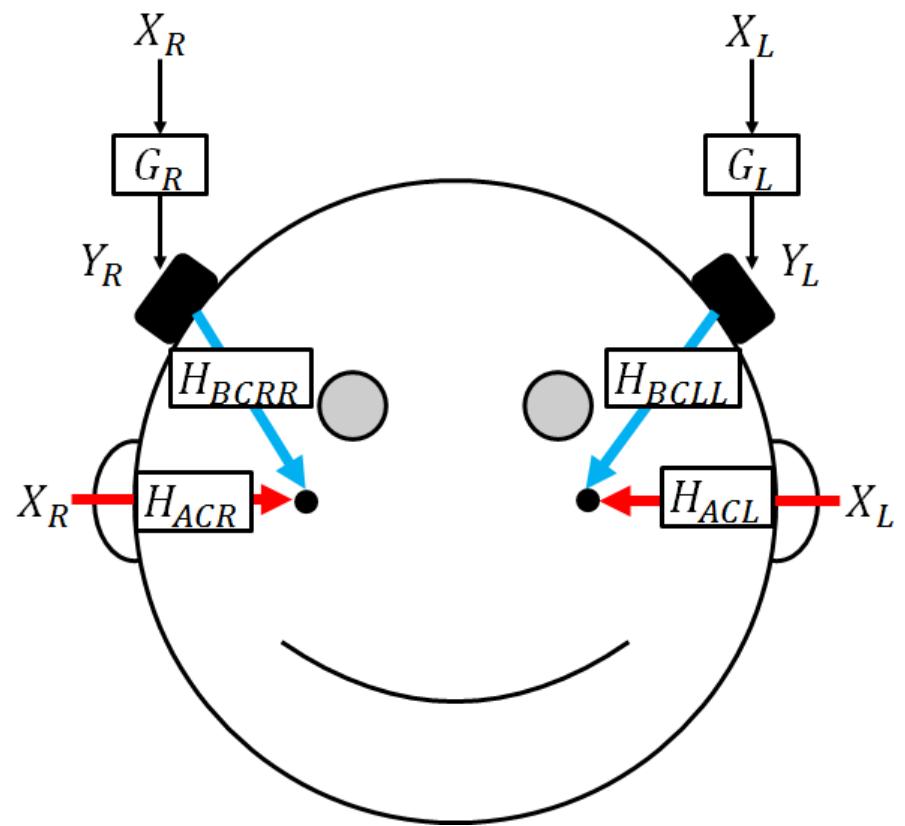


図 3.4 提案手法 1

3.3 提案手法2（「クロストーク」の除去）

3.2節で用いた前処理フィルタは、右（左）振動子から右（左）蝸牛に到達する信号を右（左）気道音と等価にするためのものであった。しかしながら、骨導音には左右の振動子から反対側の蝸牛に信号が到達してしまう「クロストーク」の問題があり、提案手法1ではこれについては考慮をしていない。

そこで、2つ目の提案手法は、骨伝導ヘッドホンとHRTFの組み合わせによる音像定位技術における、「クロストーク」の問題を解決する手法である。この手法では、図3.5に示すように4つの前処理フィルタ $G_{RR}(z)$, $G_{LL}(z)$, $G_{RL}(z)$, $G_{LR}(z)$ を用いることで、クロストークを除去する。ここで前節同様、外耳道入り口で再生される信号及び、その信号が外耳・中耳を経由し蝸牛に到達した信号をそれぞれ z 変換したものを $X_R(z)$, $X_L(z)$, $Z_{ACR}(z)$, $Z_{ACL}(z)$ とし、それらの関係は式(3.4), 式(3.5)と同様とする。

左右の振動子において再生される信号の z 変換を $Y_R(z)$, $Y_L(z)$, 骨を経由して蝸牛に到達する信号の z 変換を $Z_{BCR}(z)$, $Z_{BCL}(z)$ とすると、それぞれ以下のようなになる。

$$Z_{BCR}(z) = Y_R(z)H_{BCRR}(z) + Y_L(z)H_{BCLR}(z) \quad (3.12)$$

$$Z_{BCL}(z) = Y_R(z)H_{BCRL}(z) + Y_L(z)H_{BCLL}(z) \quad (3.13)$$

今、Rchの蝸牛に到達する信号について考える。ここで前節同様、気道音 $Z_{ACR}(z)$ と骨導音 $Z_{BCR}(z)$ を等価にすることを考えると以下のようになる。

$$\begin{aligned} Z_{BCR}(z) &= Z_{ACR}(z) \\ Y_R(z)H_{BCRR}(z) + Y_L(z)H_{BCLR}(z) &= X_R(z)H_{ACR}(z) \\ Y_R(z)\frac{H_{BCRR}(z)}{H_{ACR}(z)} + Y_L(z)\frac{H_{BCLR}(z)}{H_{ACR}(z)} &= X_R(z) \end{aligned} \quad (3.14)$$

同様にLchについて考えると、以下のようになる。

$$Y_R(z)\frac{H_{BCRL}(z)}{H_{ACL}(z)} + Y_L(z)\frac{H_{BCLL}(z)}{H_{ACL}(z)} = X_L(z) \quad (3.15)$$

また、式(3.14)及び、式(3.15)を行列式で表記すると以下のようになる。

$$\begin{bmatrix} \frac{H_{BCRR}(z)}{H_{ACR}(z)} & \frac{H_{BCLR}(z)}{H_{ACR}(z)} \\ \frac{H_{BCRL}(z)}{H_{ACL}(z)} & \frac{H_{BCLL}(z)}{H_{ACL}(z)} \end{bmatrix} \begin{bmatrix} Y_R(z) \\ Y_L(z) \end{bmatrix} = \begin{bmatrix} X_R(z) \\ X_L(z) \end{bmatrix} \quad (3.16)$$

$Y_R(z)$, $Y_L(z)$ は, $X_R(z)$, $X_L(z)$ が前処理フィルタ $G_{RR}(z)$, $G_{LL}(z)$, $G_{RL}(z)$, $G_{LR}(z)$ を通ったものであるため, 以下のように表せる.

$$Y_R(z) = X_R(z)G_{RR}(z) + X_L(z)G_{LR}(z) \quad (3.17)$$

$$Y_L(z) = X_R(z)G_{RL}(z) + X_L(z)G_{LL}(z) \quad (3.18)$$

また, 上式を行列表記すると以下のようになる.

$$\begin{bmatrix} Y_R(z) \\ Y_L(z) \end{bmatrix} = \begin{bmatrix} G_{RR}(z) & G_{LR}(z) \\ G_{RL}(z) & G_{LL}(z) \end{bmatrix} \begin{bmatrix} X_R(z) \\ X_L(z) \end{bmatrix} \quad (3.19)$$

ここで, 式 (3.16) と式 (3.19) を比較することで以下の関係を得ることができる.

$$\begin{bmatrix} G_{RR}(z) & G_{LR}(z) \\ G_{RL}(z) & G_{LL}(z) \end{bmatrix} = \begin{bmatrix} \frac{H_{BCRR}(z)}{H_{ACR}(z)} & \frac{H_{BCLR}(z)}{H_{ACR}(z)} \\ \frac{H_{BCRL}(z)}{H_{ACL}(z)} & \frac{H_{BCLL}(z)}{H_{ACL}(z)} \end{bmatrix}^{-1} \quad (3.20)$$

3.1 節において述べた推定手法により, $(H_{BCRR}(z)/H_{ACR}(z))$, $(H_{BCLR}(z)/H_{ACR}(z))$, $(H_{BCRL}(z)/H_{ACL}(z))$, $(H_{BCLL}(z)/H_{ACL}(z))$ は推定可能であり, 式 (3.20) より, クロストークを除去する前処理フィルタが作成可能である.

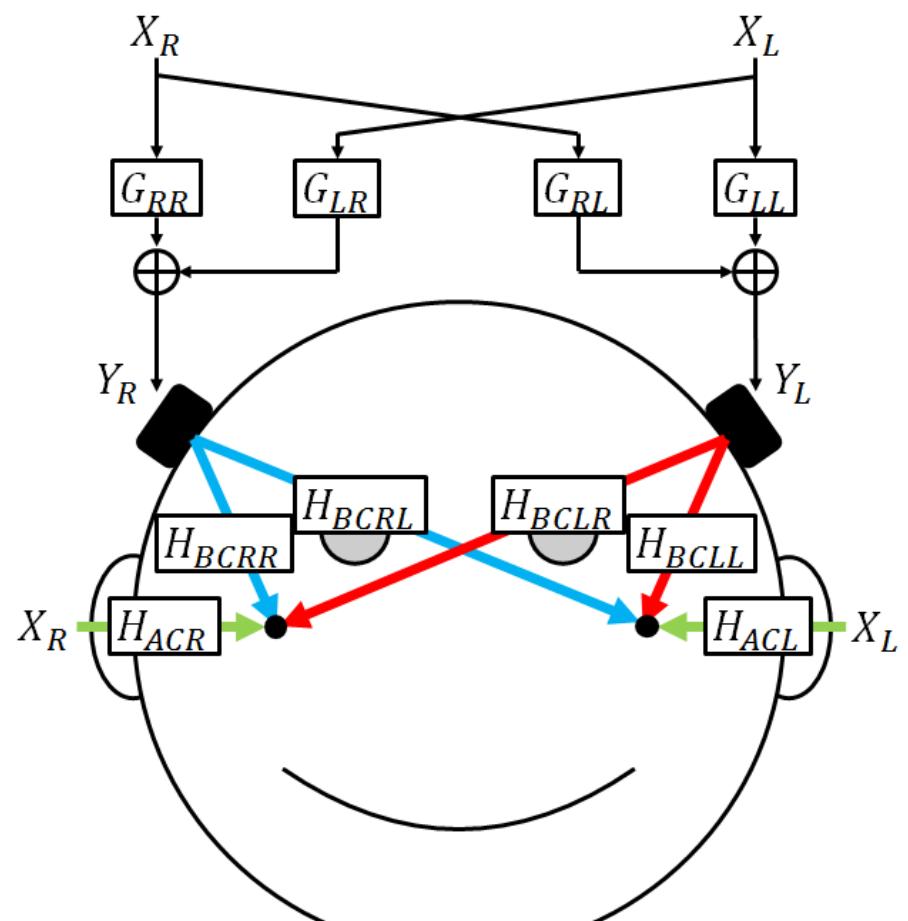


図 3.5 提案手法2

3.4 まとめ

本章では、気道音と骨導音の伝達特性差を用いた、骨伝導ヘッドホンによる音像定位精度向上のための前処理手法を提案した。提案手法1は「音の再生位置の違い」を解決する手法であり、右（左）気道音に対する右（左）骨導音の伝達特性差を用いた2つの前処理フィルタを用いる。提案手法2は「クロストーク」を除去する手法であり、4つの伝達特性差を用いた前処理フィルタを用いる。次章では、実際に被験者評価試験による音像定位試験を行い、提案手法が音像定位精度に与える影響について検証した結果を示す。

第4章 提案手法の検証

本章では、骨伝導ヘッドホンと HRTF の組み合わせによる音像定位技術に、提案手法が与える影響を被験者評価試験を通して検証する。まずははじめに、4.1節において、バイノーラル音源の測定について述べる。このバイノーラル音源は、被験者評価試験の際に用いるものであり、各被験者ごとに測定を行う。次に、4.2節において、気道音・骨導音の伝達特性差の推定条件、結果について述べる。最後に、被験者評価試験の結果を示す。

4.1 バイノーラル音源の測定条件

本研究の被験者評価試験に用いたバイノーラル音源の測定条件について述べる。音像定位の被験者評価試験において、被験者に対して様々な方位角からのバイノーラル音源を用いる。本研究では被験者の水平面上の音源に対して音像定位試験を行うため、それに対応するバイノーラル音源を測定する。バイノーラル音源の測定は、図 4.1 に示すように、被験者の水平面上に音源を設置し、被験者に対する音源の角度を変えながら 12 方向の測定を行う。測定環境は無響室であり、被験者を横から見た場合の測定の簡略図を図 4.2 に示す。また、音源から放射された信号を録音するために、マイクロホンを外耳道入り口に設置する。測定条件を以下の表 4.1 に示す。

表 4.1 バイノーラル音源の測定条件

測定環境	無響室
出力信号	白色雑音
サンプリング周波数	48kHz
信号長	2.0s
角度分解能	30 °
マイクロホン	audio-technica BP896cW
スピーカ	ECLIPSE TD307MK2A

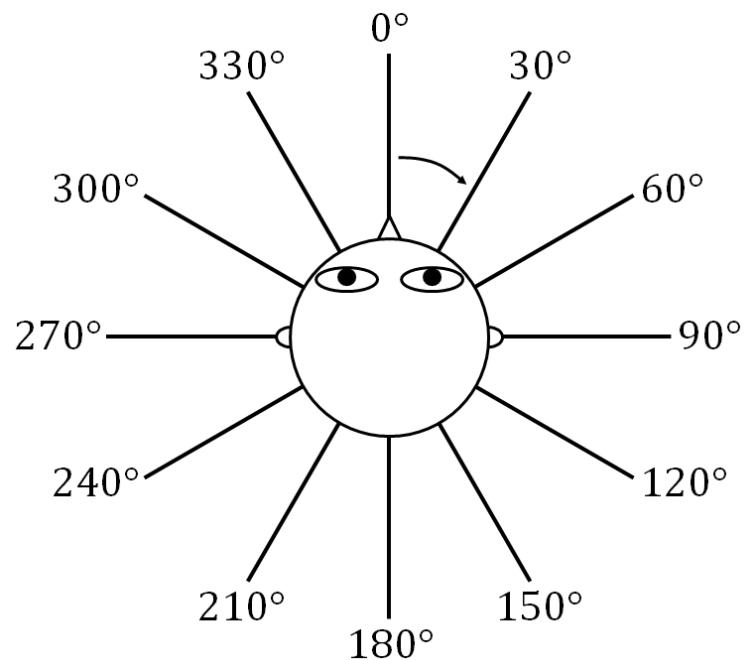


図 4.1 被験者を上から見た図

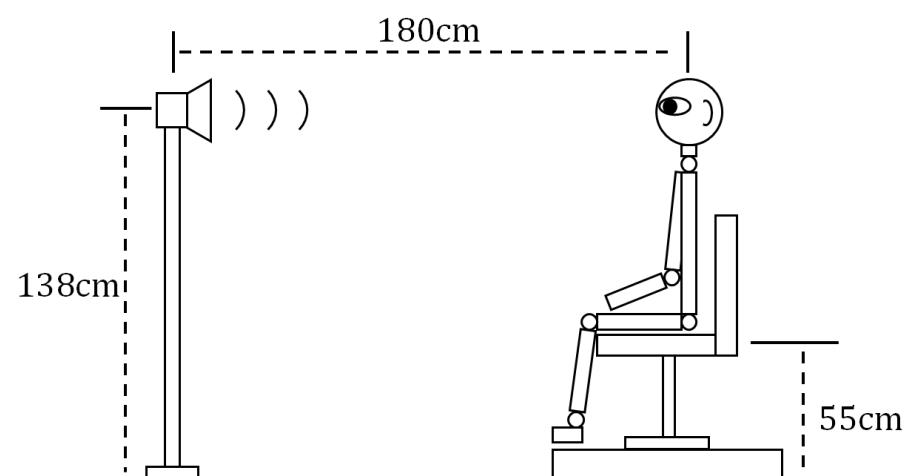


図 4.2 被験者を横から見た図

4.2 気道音・骨導音の振幅特性差の推定

本節では、2.3節で述べた気道音と骨導音の振幅特性差の推定手法を用いた、被験者ごとの推定結果を示す。推定に用いるDPOAEの測定条件を以下の表4.2に示す。DPOAE発生用の2つの刺激音の周波数をそれぞれ f_1, f_2 ($f_1 < f_2$) とし、音圧を L_1, L_2 とする。また、測定の簡略図及び、実際の測定時の写真を図4.3と図4.4に示す。

表 4.2 振幅特性差の推定条件

測定環境	無響室
被験者数	7人
サンプリング周波数	48kHz
刺激音長	100ms
f_2 範囲	1000Hz - 10000Hz
f_2 刻み	500Hz
f_2 / f_1	1.22
L_1	75dB SPL
L_2 範囲(気道)	45 - 75dB SPL
L_2 範囲(骨導)	55 - 85dB SPL
L_2 刻み	5dB SPL
刺激音出力機器(気道)	Etymotic Research ER-2
刺激音出力機器(骨導)	TEAC HP-F200
マイク	Etymotic Research ER-10B

被験者7人それぞれの、推定した振幅特性差を図4.5から図4.8に示す。図4.5に左外耳道入り口から左蝸牛までの振幅特性に対する、左振動子から左蝸牛までの振幅特性の差を示す。図4.6に右外耳道入り口から右蝸牛までの振幅特性に対する、右振動子から右蝸牛までの振幅特性の差を示す。図4.7に左外耳道入り口から左蝸牛までの振幅特性に対する、右振動子から左蝸牛までの振幅特性の差を示す。図4.8に右外耳道入り口から右蝸牛までの振幅特性に対する、左振動子から右蝸牛までの振幅特性の差を示す。

図4.5から図4.8に示した、各被験者の気道音と骨導音の振幅特性差より、それぞれの振幅特性差において個人差が存在することが分かる。これは、ヒトによって異なる頭蓋骨の形状や骨の密度などが、骨伝導による音の受聴に影響を与えていていることを示している。また、図4.5と図4.7、図4.6と図4.8より、振動子と同じ側にある蝸牛に伝わる音に比べて、クロストーク成分は減衰して蝸牛に伝わっていることが分かる。

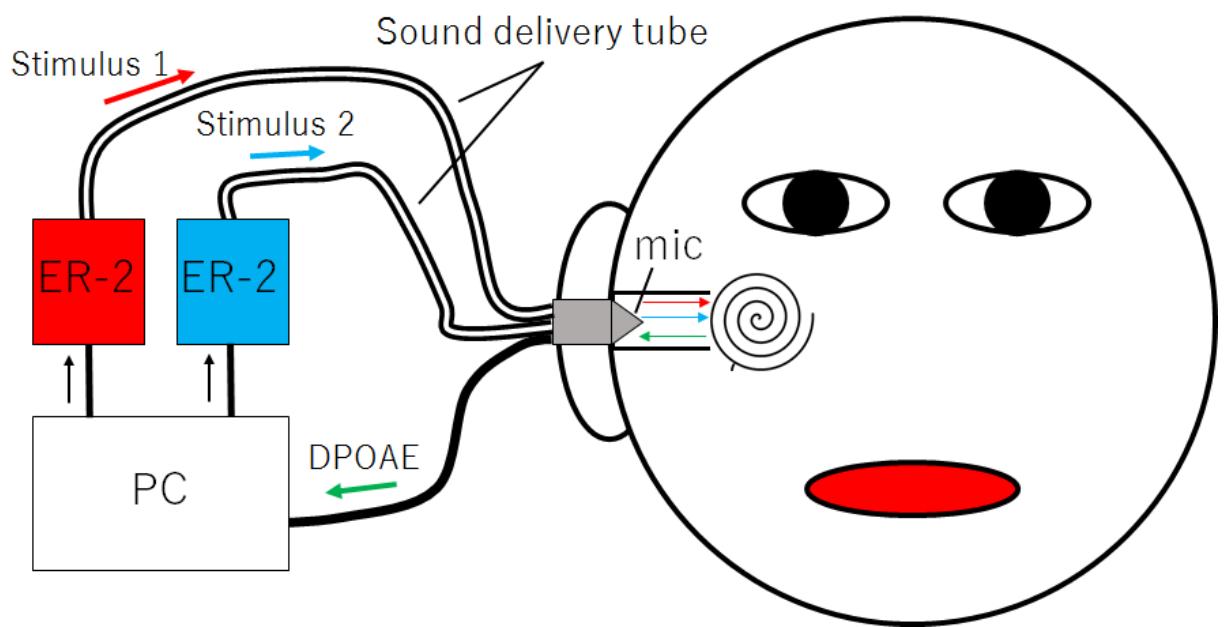


図 4.3 DPOAE 測定の簡略図



図 4.4 測定風景

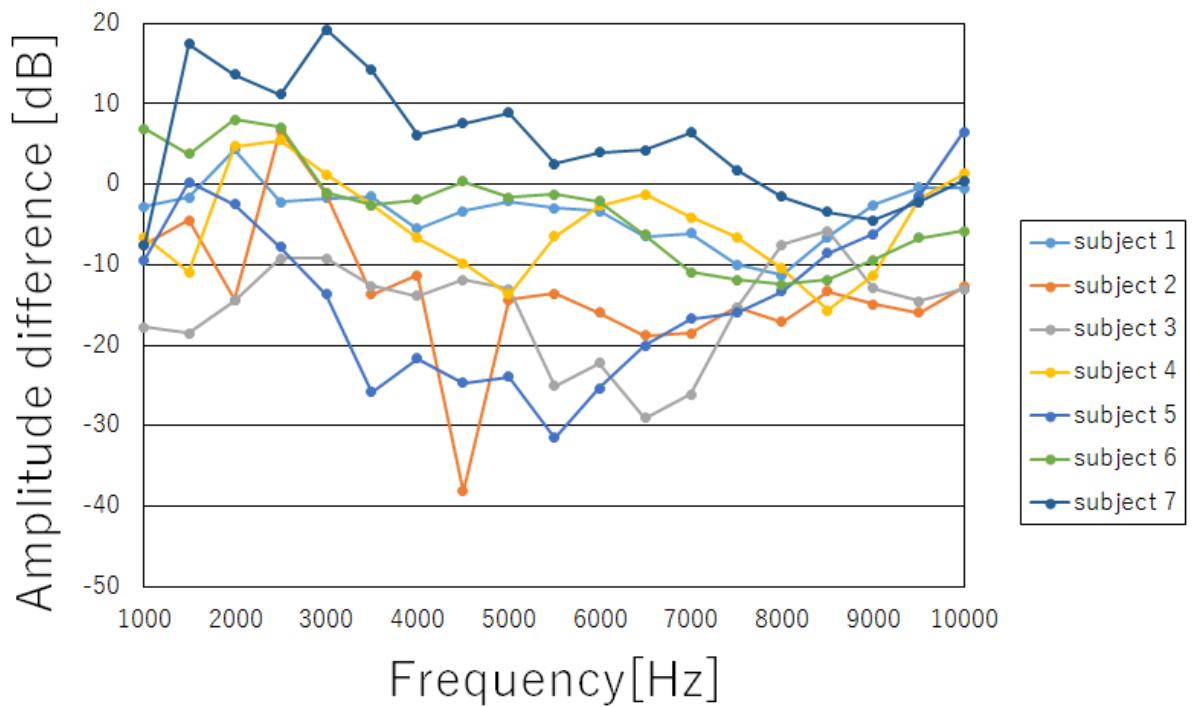


図 4.5 左気道音・左骨導音

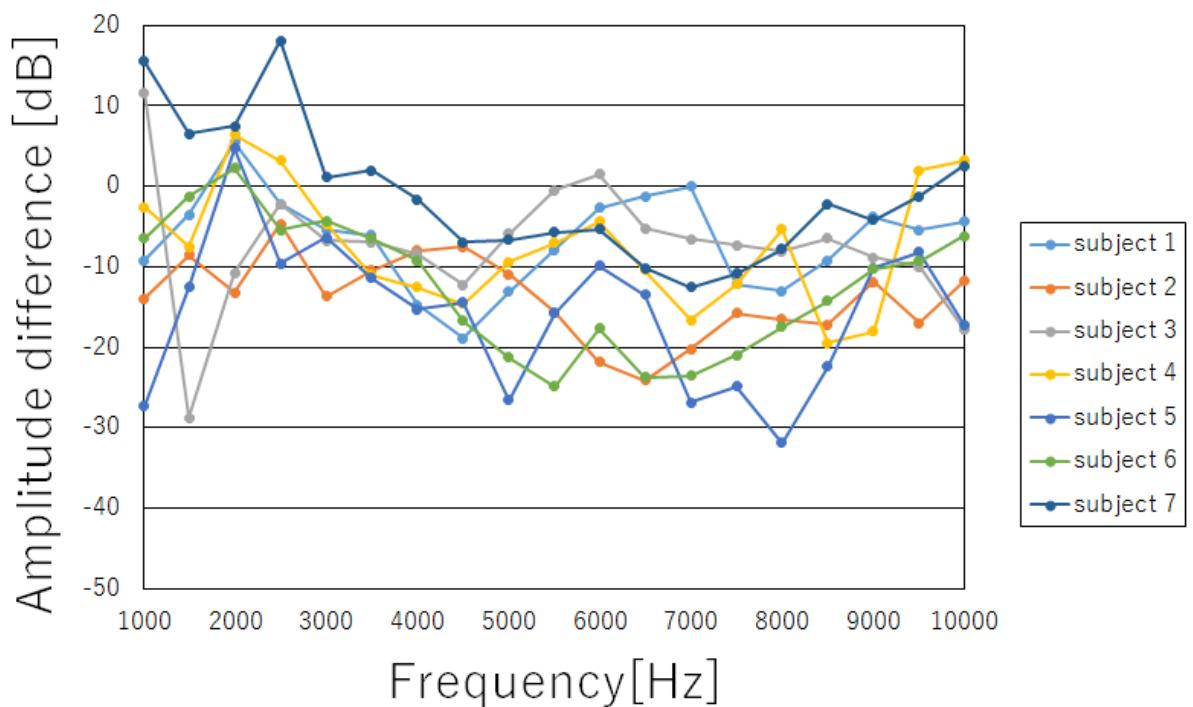


図 4.6 右気道音・右骨導音

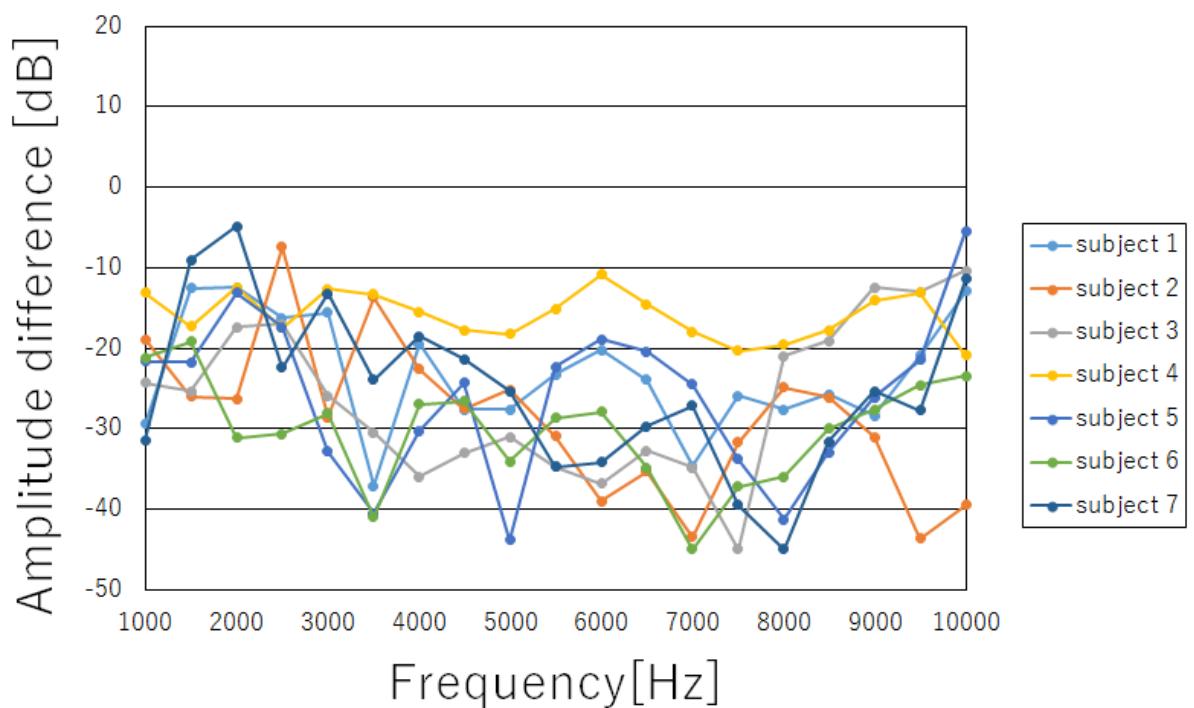


図 4.7 左気道音・右骨導音

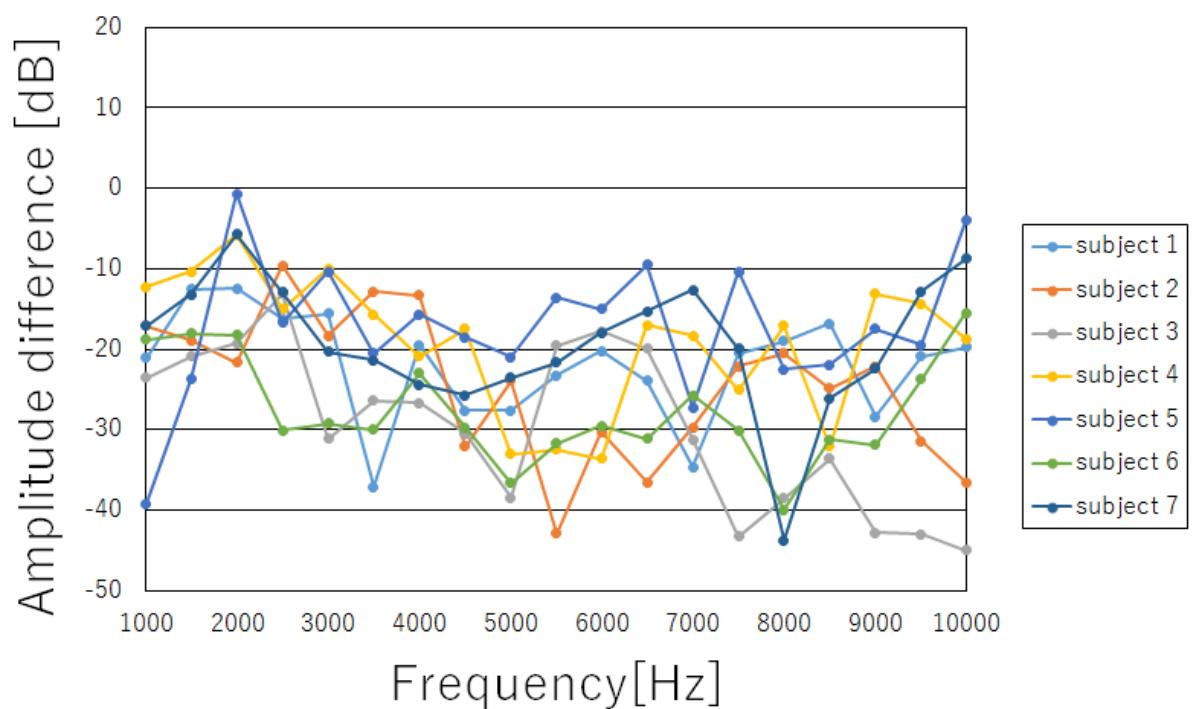


図 4.8 右気道音・左骨導音

4.3 被験者評価試験

前節の伝達特性差を用いて、3.2節の提案手法1及び、3.3節の提案手法2の前処理フィルタを作成し、音像定位の被験者評価試験を行った。提案手法1・提案手法2の前処理フィルタを作成するために、3.1節で述べたように、前節で推定した振幅特性差を実現するフィルタを作成した。作成にあたり、右（左）振動子から右（左）蝸牛までと、左（右）振動子から右（左）蝸牛までの信号の伝達時間差を考慮する必要がある。しかし、骨導音の伝播速度は4080m/s程度[15]であり、片方の振動子から左右それぞれの蝸牛までの距離差を5cmと仮定すると、到達時間差は約0.012msとなる。今回の検証ではサンプリング周波数は48kHzで行っており、サンプル差は1sampleに満たないため、伝達特性差を実現する全てのフィルタを256次の線形位相フィルタとして作成した。

音像定位試験は、4.1節で述べた被験者ごとのバイノーラル音源を用いて、気道音、前処理無しの骨導音、提案手法1,2を適用した骨導音の4種類の提示方法について行った。被験者にはランダムで提示される角度ごとのバイノーラル音源を受聴し、それぞれで知覚した方向を回答させた。今回バイノーラル音源はサンプリング周波数48kHzで録音したが、振幅特性差は10000Hzまでの推定であったため、LPFによりバイノーラル音源を周波数10000Hzまでに帯域制限した。

以下の表に音像定位試験の条件を示す。

表 4.3 音像定位試験条件

被験者数	7人
試験信号	バイノーラル音源
提示回数	各方向3回
提示方法	ランダム
提案手法1の各フィルタ次数	513
提案手法2の各フィルタ次数	1281
イヤホン	Roland CS-10EM
骨伝導ヘッドホン	TEAC HP-F200

音像定位試験の結果を以下の図4.9から図4.12に示す。このグラフは、ある角度のバイノーラル音源を受聴した際の、全被験者の回答した角度の合計回数を表している。グラフの横軸は録音したバイノーラル音源の被験者に対する角度方向となっており、縦軸は被験者が知覚した音源の方向を表している。グラフ上の円は、回答回数を大きさで表している。

すなわち，仮に全被験者が全てのバイノーラル音源に対して，正しい方向を知覚・回答した場合には，グラフの左下から右上にかけての対角線上に大きな円が並ぶことになる。平均正答率を以下の表 4.4 に示す。今回の試験では全被験者が提案手法 1 を適用した場合には，前処理無しの骨伝導ヘッドホンを用いた場合よりも定位精度が向上した。また，提案手法 2 を適用した場合には，全被験者の定位精度は低下した。

表 4.4 音像定位精度

気道音	骨導音	提案手法 1	提案手法 2
70.6%	32.5%	49.2%	21.0%

図 4.9 と図 4.10 を比較すると，気道音による試験では高い音像定位精度が実現できている。それに対して，骨導音による試験では前後の不一致や，音源方向の両隣の角度に知覚している場合が多く，定位精度は低い。図 4.10 と，提案手法の結果である図 4.11 と図 4.12 を比較する。提案手法 1 を適用することで，前処理無しの骨伝導ヘッドホンを用いた場合よりも，音源方向に対する知覚のずれが少くなり，定位精度が向上していることが分かる。しかし提案手法 2 を適用した場合には，前処理無しの骨伝導ヘッドホンよりも，音源方向に対する知覚のずれが大きくなり，定位精度が下がってしまっていることが分かる。

提案手法 1 によって定位精度が上がったことから，気道音・骨導音の振幅特性差による補正は，定位精度向上に効果的だと考えられる。しかし，提案手法 2 によって定位精度が下がってしまったことから，定位精度低下の要因として「クロストーク」の除去ができないことが考えられる。クロストークは左右の振動子から反対側の蝸牛に到達する信号であるため，3.3 節において示したように，提案手法 2 では 4 つのフィルタを用いて右（左）の信号を左（右）からも出力することでクロストークの除去を行っている。しかし，この手法によってクロストークの除去ができない場合には，左右の信号が混ざった信号を受聴することになる。そのため，定位精度が低下していると考えられる。クロストークが除去できていない要因として，以下のことが考えられる。今回，骨導音の伝播速度が非常に速いため，左右の振動子から蝸牛までの到達時間が同じと仮定して，伝達特性差のフィルタを全て同じ次数として作成した。しかし，クロストークの除去ができないことから，実際の骨導音の伝播速度は今回用いたもの [15] よりも遅く，左右の振動子から蝸牛までの到達時間差を考慮して伝達特性差のフィルタを作成する必要があると考えられる。よって，提案手法 2 において，クロストークの除去を実現するには，左右の振動子から蝸牛までの

到達時間差の推定が必要だと考えられる。

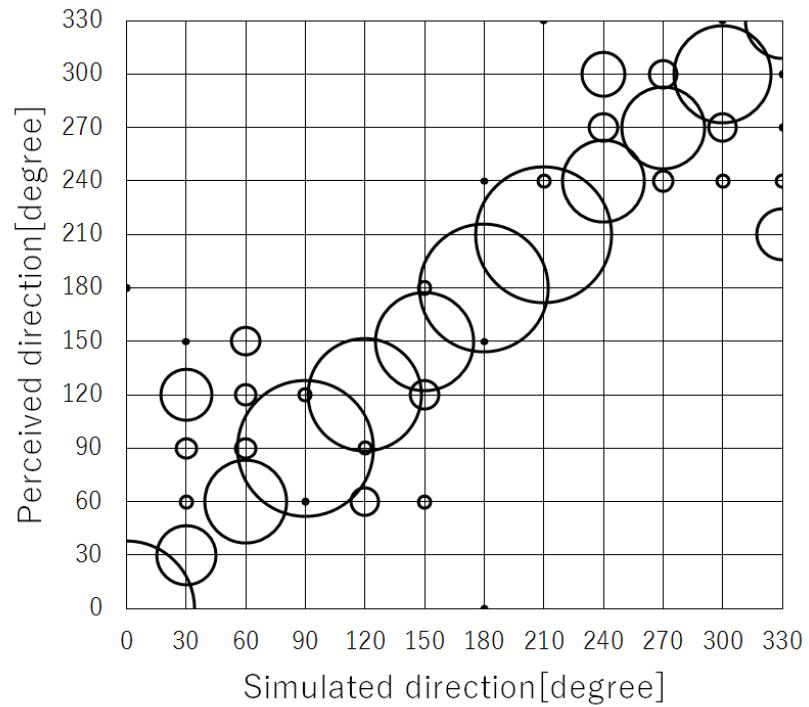


図 4.9 気道音による試験結果

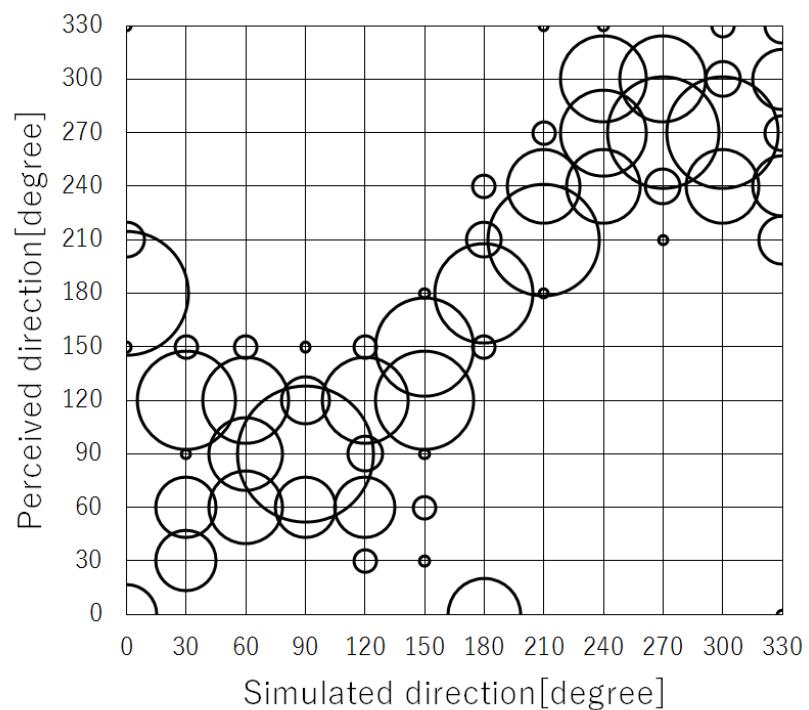


図 4.10 前処理無しの骨導音による試験結果

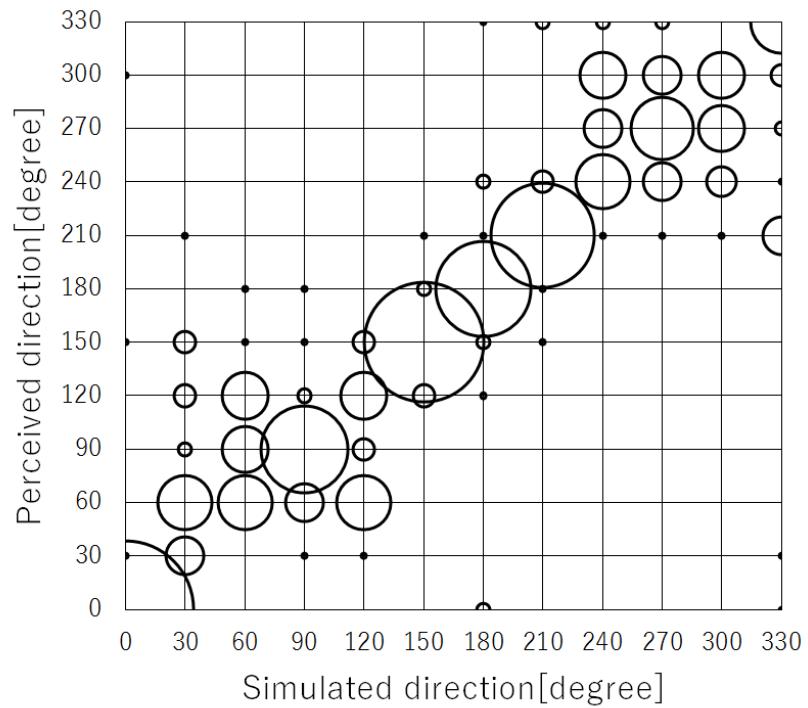


図 4.11 提案手法 1 による試験結果

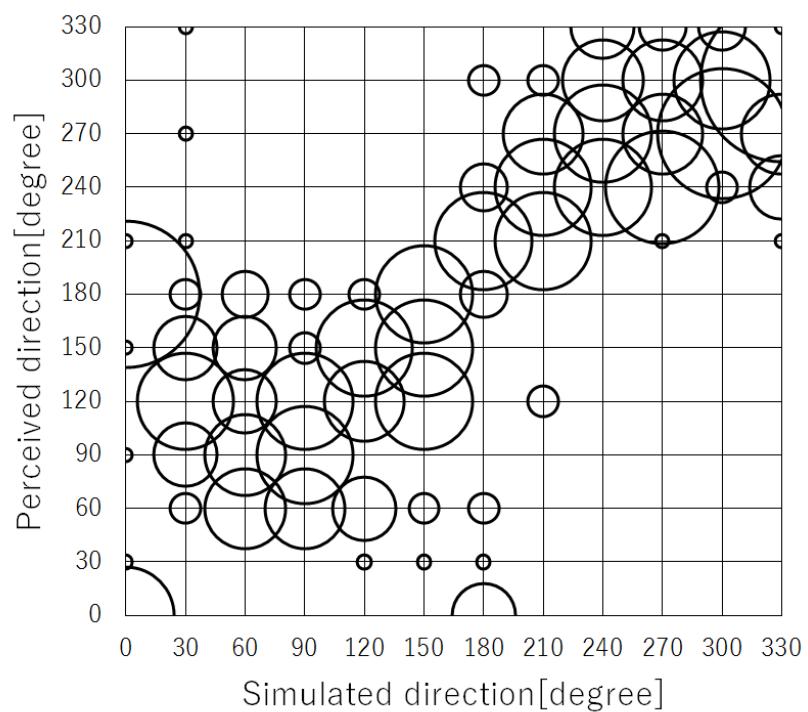


図 4.12 提案手法 2 による試験結果

4.4 まとめ

本章では、音像定位精度の被験者評価試験を行い、2つの提案手法による音像定位精度への影響を検証した。振幅特性差の推定において、被験者間の振幅特性差には違いがあることを示した。被験者評価試験では、被験者ごとに水平面上 30° 刻みで録音したバイノーラル音源を用いて音像定位試験を行った。その結果、提案手法1を適用した場合には、全ての被験者の定位精度が前処理無しの骨伝導ヘッドホンを用いた場合の定位精度よりも向上した。しかし、提案手法2を適用した場合には、全ての被験者の定位精度が下がってしまった。この原因として、クロストークの除去ができないことが挙げられ、これを解決するために、左右の振動子から蝸牛までの到達時間差の推定を行う必要があると考えられる。また、提案手法1によって定位精度が上がったことから、振幅特性差の推定手法は音像定位精度向上に効果的であるが、被験者1人にかかる測定時間が長いため、測定時間を短縮することが望ましい。

第 5 章 おわりに

本論文では、骨伝導ヘッドホンと HRTF の組み合わせによる音像定位精度向上を目的とし、「音の再生位置の違い」と「クロストーク」による問題点を解決する前処理手法を提案した。第 1 章では、骨伝導ヘッドホンによる音像定位精度の利点とその現状及び、定位精度低下の要因について述べ、本論文の研究目的について述べた。第 2 章では、本論文における基礎的な知識について説明し、本研究の基礎となる、外耳道入り口から蝸牛までの振幅特性に対する振動子から蝸牛までの振幅特性の差の推定手法について説明した。第 3 章では、本論文の提案手法である、伝達特性差を用いた前処理手法について述べた。提案手法 1 は「音の再生位置の違い」を解決する手法であり、2 つの前処理フィルタで構成される。提案手法 2 は「クロストーク」を解決する手法であり、4 つの前処理フィルタで構成される。第 4 章では、被験者評価試験を通して、提案手法が音像定位精度に与える影響について検証した。その結果、全ての被験者において提案手法 1 を適用した場合には、前処理無しの骨伝導ヘッドホンの場合よりも定位精度が向上した。このことから、振幅特性差の推定手法を用いる前処理手法は、音像定位精度向上に効果的であることを示した。しかし、提案手法 2 を適用した場合には、全ての被験者の定位精度が低下してしまった。その要因として、クロストークの除去ができていないことが考えられ、この問題を解決するために、左右の振動子から蝸牛までの信号の到達時間差を推定し考慮する必要があると考えられる。

謝辞

本研究を進めるにあたり、日頃より丁寧なご指導ご鞭撻を頂きました、本学電気系准教授杉田泰則博士に、深謝いたします。また、本論文の審査において貴重なご意見を賜りました、本学電気系教授岩橋政宏博士、並びに本学電気系准教授武井由智博士に御礼申し上げます。さらに、本研究の被験者を快く引き受けてください、そして多くのご指摘を下さった、杉田研究室の同期・後輩の皆様に感謝の意を表します。

平成 29 年 2 月

参考文献

- [1] T. Miura, Y. Ebihara, M. Sakajiri, T. Ifukube, "Utilization of auditory perceptions of sounds and silent objects for orientation and mobility by visually-impaired people," IEEE International Conference on Systems, Man, and Cybernetics, pp.1080-1087, 2011.
- [2] Y. Seki, T. Sato, "A training system of orientation and mobility for blind people using acoustic virtual reality," IEEE Transaction on Neural Systems Rehabilitation Engineering, vol.19, no.1, pp.95-104, 2011.
- [3] 松崎成敏, 黒田嘉宏, 浦西友樹, 井村誠孝, 大城理, "立体音響のための骨伝導の信号伝播解析," 電子情報通信学会技術研究報告, Vol. 113(241), pp. 25-30, 2013.
- [4] 浅野雄二, 村山淳, 相川直幸, 原田哲也. "骨伝導ヘッドホンを用いた立体音響提示のためのクロストークキャンセルに関する一検討," 電子情報通信学会技術研究報告, Vol. 111(404), pp. 127-132, 2013.
- [5] Stefan Stenfelt, Richard L. Goode, "Transmission properties of bone conducted sound: Measurements in cadaver heads," The Journal of the Acoustical Society of America, Vol. 118(4), pp. 2373-2391, 2005.
- [6] Stefan Stenfelt, Bo Hakansson, Anders Tjellstrom, "Vibration characteristics of bone conducted sound in vitro," The Journal of the Acoustical Society of America, Vol. 107(1), pp. 422-431, 2000.
- [7] 町田章一, 嶋倉優子, 岡本途也, 野々村英一, "骨導補聴器の研究," 日本聴覚医学会, Vol. 26, pp. 27-33, 1983.
- [8] D. T. Kemp, "Stimulated acoustic emissions from within the human auditory system," The Journal of the Acoustical Society of America, Vol. 64(5), pp. 1386-1391, 1978.
- [9] Hans Kunov, David W. Purcell, William Cleghorn "Estimating bone conduction transfer functions using otoacoustic emissions," The Journal of the Acoustical Society of America, Vol. 114(2), pp. 907-918, 2003.

- [10] 飯田一博, 森本政之, “空間音響学,” コロナ社, 2010.
- [11] M. Morimoto, Y. Ando “On the simulation of sound localization,” Journal of the Acoustical Society of Japan, Vol. 1(3), pp. 167-174, 1980.
- [12] F. P. Harris, B. L. Lonsbury-Martin, B. Stagner, A. C. Coats, G. K. Martin, “Acoustic distortion products in humans: Systematic changes in amplitude as a function of f_2/f_1 ratio,” The Journal of the Acoustical Society of America, Vol. 85(1), pp. 220-229, 1989.
- [13] B. L. Lonsbury-Martin, F. P. Harris, B. Stagner, M. D. Hawkins, G. K. Martin, “Distortion product emissions in humans. I. Basic properties in normally hearing subjects,” Ann Otol Rhinol Laryngol, Vol. 147, pp. 3-14, 1990.
- [14] M. L. Whitehead, B. Stagner, M. J. McCoy, B. L. Lonsbury-Martin, G. K. Martin, “Dependence of distortion-product otoacoustic emissions on primary levels in normal and impaired ears. I. Effects of decreasing L2 below L1,” The Journal of the Acoustical Society of America, Vol. 97(4), pp. 2346-2358, 1995.
- [15] 菅原勇勝, “柔道整復治療法,” ヘルスシステム研究所, 2004.