

長岡技術科学大学大学院

工学研究科修士論文

題 目

骨導立体音像定位のための  
伝達特性の推定に関する研究

指導教員            准教授   杉田 泰則

著者   電気電子情報工学専攻  
13110090 横田 佑香里

提出期日            平成31年2月8日

# 目次

<b>第 1 章</b>	<b>はじめに</b>	<b>1</b>
1.1	研究背景 . . . . .	1
1.2	研究目的 . . . . .	1
1.3	本論文の構成 . . . . .	2
<b>第 2 章</b>	<b>本研究の基礎となる知識・手法</b>	<b>3</b>
2.1	頭部伝達関数 [9] . . . . .	3
2.2	歪み成分耳音響放射 (Distortion Product OtoAcoustic Emis- sion, DPOAE) [8] . . . . .	6
2.3	振幅特性差の推定手法 [17] . . . . .	7
2.4	骨導音による立体音像定位手法 [7] . . . . .	10
2.5	スパースモデル [18, 19] . . . . .	11
2.6	まとめ . . . . .	12
<b>第 3 章</b>	<b>HRTF 合成手法</b>	<b>13</b>
3.1	主成分分析を用いた HRTF 合成 [13] . . . . .	13
3.2	提案法 . . . . .	14
3.2.1	提案法 1-1 (PCA+スパース) . . . . .	14
3.2.2	提案法 1-2(誤差推定) . . . . .	15
3.3	実験 . . . . .	16
3.4	まとめ . . . . .	20
<b>第 4 章</b>	<b>DPOAE 抽出手法</b>	<b>21</b>
4.1	DPOAE 抽出手法 (DFT) . . . . .	21
4.2	提案法 2-1(スパース推定) . . . . .	21
4.3	実験 . . . . .	24

4.3.1	DPOAE 疑似データ . . . . .	24
4.3.2	DPOAE 実測データ . . . . .	28
4.4	まとめ . . . . .	31
第 5 章 おわりに		32
謝辞		33
付録 A 付録		34
A.1	付録 1 . . . . .	34
A.2	付録 2 . . . . .	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．鼓膜を介して音を受聴する一般的な気道ヘッドホンに関しては古くから研究が行われており，利用者個人の頭部伝達関数（Head-Related Transfer Function, HRTF）を用いることで高い音像定位精度を実現できる [1]．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，耳介などの影響によって生じる音圧変化を表すものである．

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盛んに行われており [2–4]，映画やゲームなどのエンターテインメント分野のほか，視覚障害者向けの歩行支援機器などの医療福祉分野への応用も期待されている [5, 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．気道ヘッドホンと異なり耳を塞がないため，周囲の環境音を受聴することが可能である．

気道ヘッドホンの場合は，利用者個人の 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．一方で，骨伝導ヘッドホンの場合は，利用者個人の HRTF を用いても音像定位精度が低いといった問題がある．この問題は，気道ヘッドホンとの「音の再生位置の違い」に起因するものである．多田らは，外耳道入り口から蝸牛まで（気道）と，振動子から蝸牛まで（骨導）の伝達特性差を前処理として利用する手法を提案しており，これにより再生位置の違いを補正することができ，音像定位精度が向上することを示している [7]．

骨導立体音像定位における伝達特性差の推定には，生体反応である歪み成分耳音響放射（Distortion Product OtoAcoustic Emission, DPOAE）を用いている [8]．しかし，HRTF と DPOAE の測定には多大な労力と時間を要し，利用者の負担が大きいという問題がある．

## 1.2 研究目的

本論文では，測定時間の短時間化を目的とし，主成分分析による基底展開とスパースモデルを用いて部分的に測定した HRTF から未測定方向の HRTF を合成する手法を提案し，従来法よりも高い合成精度を目指す．また，スパース推定を用いて短い測定時間から伝達特性差の推定に必要な DPOAE を抽出する手法を提案し，従来法では抽出ができなかった DPOAE の抽出を目指す．

### 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．第3章では，従来法及び提案法による未測定方向のHRTFを合成する手法について述べ，合成精度に関して従来法と比較する．第4章では，従来法及び提案法によるDPOAEの抽出手法について述べ，疑似データと実測データを用いて従来法と抽出精度を比較し，有用性を検証する．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、本論文で用いる基礎的な知識・手法について述べる．2.1 節では、頭部伝達関数を用いた立体音像定位手法について述べる．2.2 節では、生体反応である歪み成分耳音響放射について述べる．2.3 節では、歪み成分耳音響放射を用いた気道と骨導の振幅特性差の推定手法について述べる．2.4 節では、振幅特性差を用いた骨導音による立体音響定位手法について述べる．2.5 節では、提案法に用いるスパースモデルについて述べる．

### 2.1 頭部伝達関数 [9]

本節では、音像定位、座標系及び頭部伝達関数について説明する．

始めに、音像の概念を説明する．音源から発せられた音波が鼓膜に到達した際に、知覚したものの総体を音像といい、音像の方向と距離を知覚することを音像定位という．また、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のことである．気道ヘッドホンに関しては、頭部伝達関数（Head-Related Transfer Function, HRTF）を用いる手法が一般的である．

次に、本論文で用いる座標系を説明する．図 2.1 に音源の方向を表すための座標系を示す．本論文では頭を中心とした球座標系を用いる．この座標系は次のように定義される．原点は左右の外耳道を結ぶ線分の midpoint である．図 2.2 に示すように、水平面は眼窩点（眼球の収まる頭蓋骨のくぼみ）と耳珠（外耳道入り口のでっぱり）を含む平面であり、横断面は左右のが外耳道入り口を通り水平面に直交する面である．また、正中面は水平面と横断面の両方に直交する面である．角度は、方位角  $\theta$  と仰角  $\varphi$  を用いる．本論文では、音源が水平面上に存在する場合に対しての検証を行っているため、仰角  $\varphi = 0^\circ$  はであり、以降仰角  $\varphi$  は省略する．

次に、頭部伝達関数の定義を説明する．ある音波が耳に到達するときに、受聴者の頭や耳介の形状、あるいは胴体の影響を受ける．頭部伝達関数 (Head-Related Transfer Function, HRTF) とは、このような頭部周辺の影響によって生じる入射音の音圧変化を周波数の関数で表したものをいい、頭部伝達関数  $H_{lr}(s, \theta, \omega)$  は次式で定義される．

$$H_{lr}(s, \theta, \omega) = \frac{G_{lr}(s, \theta, \omega)}{F(s, \theta, \omega)} \quad (2.1)$$

ここで、 $G_{lr}$  は反射や回折のない自由音場における音源から受聴者の左右の外耳道入り口

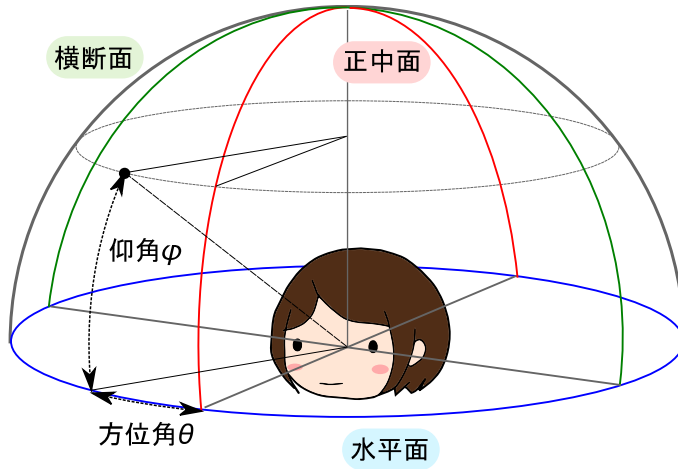


図 2.1 極座標系

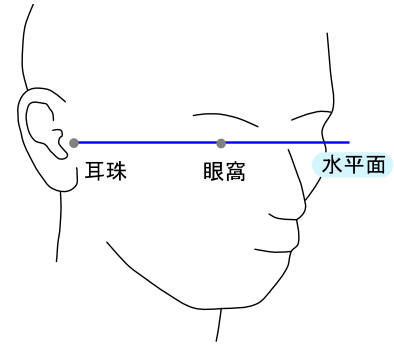


図 2.2 極座標系（水平面）

までの伝達関数であり， $F$  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．添え字  $l, r$  は耳の左右， $s$  は受聴者， $\theta$  は方位角， $\omega$  は周波数をそれぞれ表す．

図 2.3 は同一被験者に対して，正面（方位角  $\theta = 0^\circ$ ）から右（ $\theta = 90^\circ$ ）までを 30 度間隔で測定した右耳の HRTF 振幅特性である．図 2.3 のように，音源の角度によって HRTF 振幅特性が異なり，その特性の違いから受聴者は音源の方向を知覚している．図 2.4 は同一方向正面（方位角  $\theta = 0^\circ$ ）に対する，3 名の被験者の HRTF 振幅特性である．図 2.4 のように，HRTF は音源の角度が同じ場合でも個人差が生じる．他者の HRTF を用いて音像定位を行う場合，前後方向について誤判定が頻繁に発生するなど，音像定位精度が低下してしまう [10]．そのため，高い音像定位精度を得るためには，受聴者本人の HRTF を測定する必要がある．

受聴者の HRTF 測定の負担軽減のため，部分的に測定した HRTF から，未測定方向の HRTF を推定する手法が提案されている．これまで，線形補間 [11, 12] や主成分分析 [13] を用いた手法が提案されている．

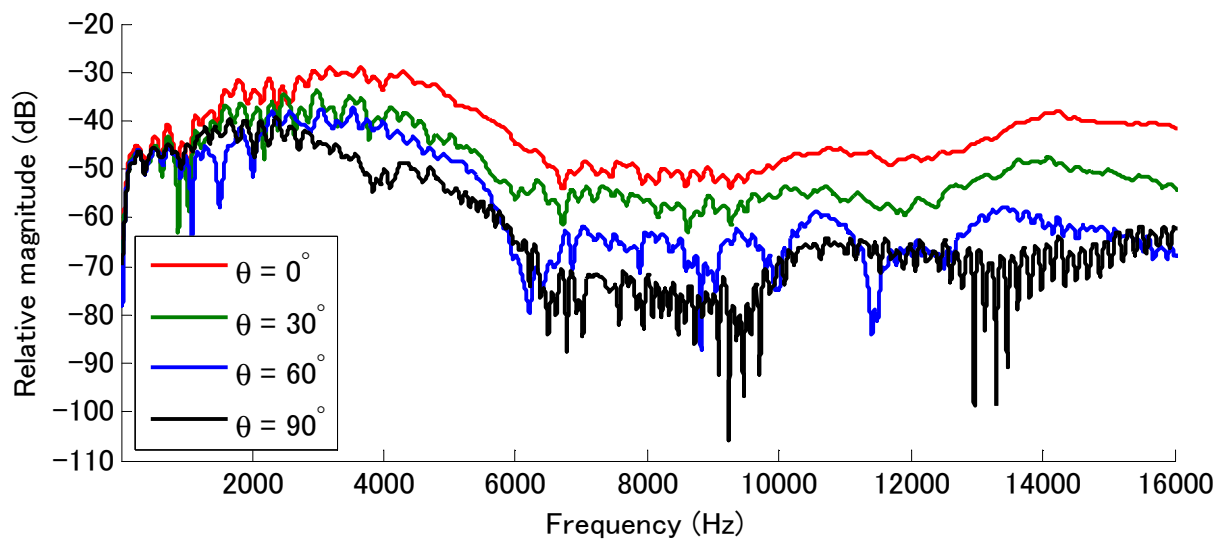


図 2.3 HRTF の角度比較例 (同一被験者)

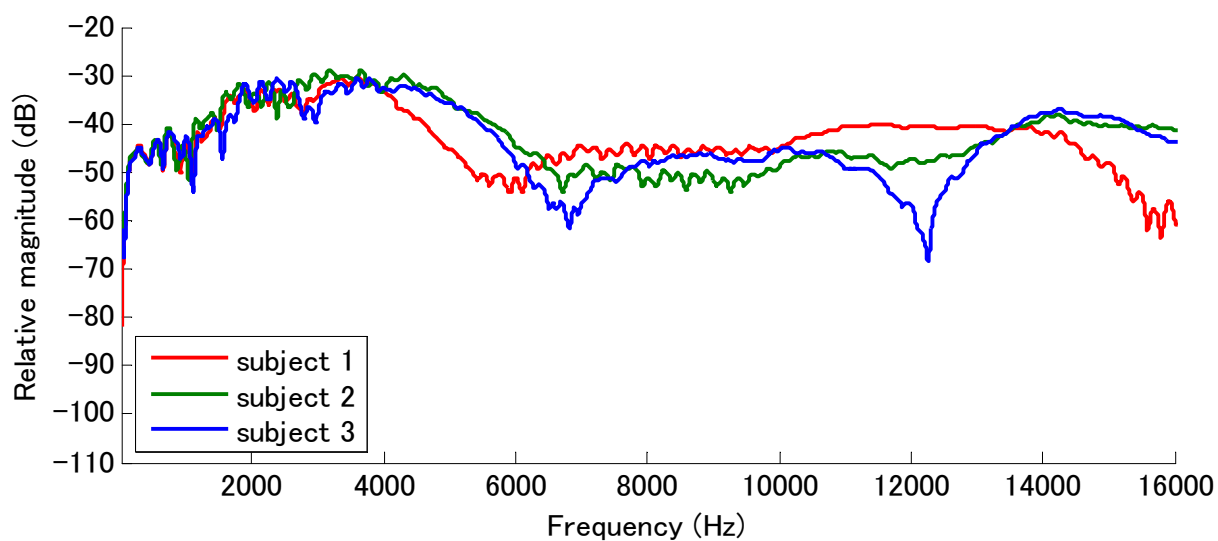


図 2.4 HRTF の被験者比較例 (方位角  $\theta = 0^\circ$ )



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

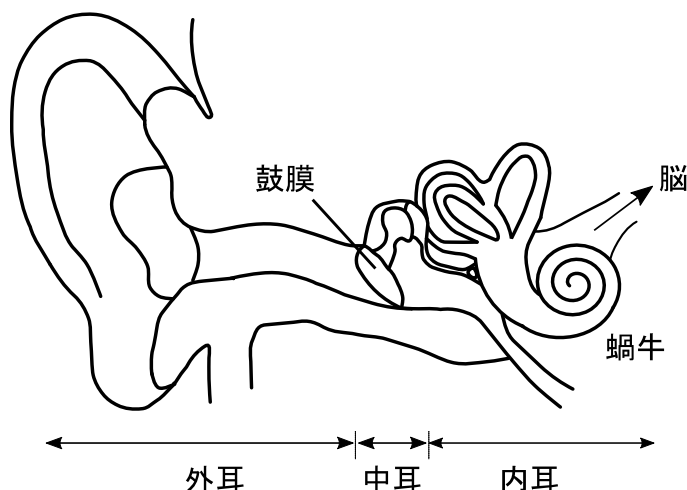


図 2.5 耳の構造

本節では、気道と骨導の振幅特性差の推定に用いる歪み成分耳音響放射について説明する．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 2.5 にヒトの耳の構造を示す．通常ヒトが音を受聴する際、外耳に入った信号は鼓膜を介し、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．これに対して、耳音響放射 (OtoAcoustic Emission, 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．

OAE は、音刺激に関係なく外耳道に放射される自発耳音響放射 (Spontaneous Otoacoustic Emission, SOAE) と、音刺激によって誘発される耳音響放射 (Evoked Otoacoustic Emission, EOAE) がある．EOAE の一種である DPOAE 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刺激音の歪成分が生成される現象である．DPOAE の信号は、刺激音の 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に由来することから、蝸牛機能を評価するのに用いられている．

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いる．周波数比を  $f_2/f_1 = 1.22$  付近で与えた

際に放射される 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 [14] . DPOAE の音圧は, 刺激音の周波数と音圧によって非線形的に変化する [15] . これらの特性から, DPOAE は聴力検査や生体認証 [16] などに用いられている.

### 2.3 振幅特性差の推定手法 [17]

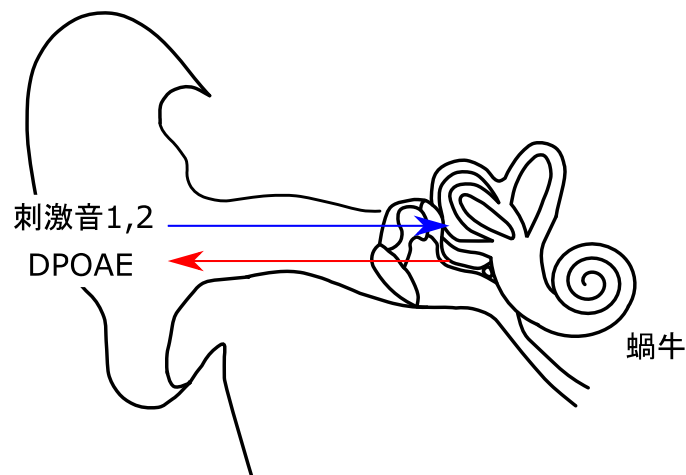


図 2.6 DPOAE 測定図 (気道)

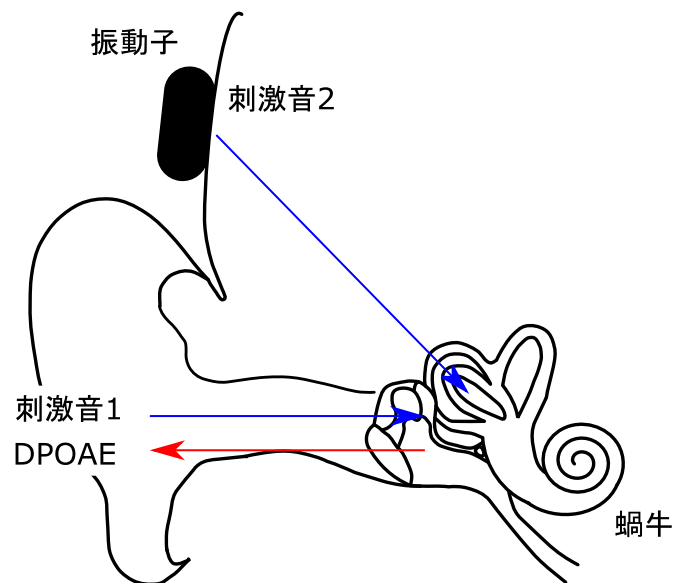


図 2.7 DPOAE 測定図 (骨導)

2.2節で述べた DPOAE を用いて、外耳道入り口から蝸牛までの伝達特性（気道）に対する、骨伝導振動子から蝸牛までの振幅特性（骨導）の差を推定する手法を説明する．2つの刺激音を気道音で与えた場合の DPOAE(図 2.6) と，気道音・骨導音それぞれで与えた場合の DPOAE(図 2.7) を比較することで，振幅特性の差を推定する．ここで，DPOAE 誘発用の 2つの刺激音の周波数をそれぞれ  $f_1, f_2 (f_1 < f_2)$ ，音圧を  $L_1, L_2$  とする．

はじめに，刺激音 1, 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， $L_2$  の値を変化させて測定を行う．刺激音の音圧が変化することで DPOAE の音圧も変化し，図 2.8 の赤線に示すようなグラフを得ることができる．図 2.8 の縦軸は DPOAE の音圧（ $2f_1 - f_2$  周波数成分）であり，横軸は刺激音 2 の音圧  $L_2$  となっている．図 2.8 の例では，測定条件は  $(f_1, f_2, L_1) = (4918\text{Hz}, 6000\text{Hz}, 75\text{dB})$  であり， $L_2$  は 45dB から 80dB まで 5dB 間隔で変化させている．次に，刺激音 1 を気道音，刺激音 2 を骨導音で与え，上記と同様の測定を行う．測定結果を図 2.8 に青線で示す． $L_2$  は 55dB から 80dB の範囲であり，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．片方のグラフを横軸方向にシフトさせ，2つのグラフの誤差が最小になるシフト量を算出する．ただし，図 2.8 の 2つのグラフは 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．図 2.9 に 2つのグラフの誤差が最小になるようにシフトした結果を示す．この例では，刺激音 2 を骨導音で与えたときのグラフを -5.0dB シフトした際に，誤差が最小値になった．つまり， $f_2=6000\text{Hz}$  において，骨導音は気道音に対して 5.0dB 減衰して蝸牛に到達するといえる．これを  $f_2$  を変化させ，各周波数において上記測定を行うことで，骨導音に対する気道音の振幅特性差を推定できる．

この伝達特性差を用いた，骨導音による立体音像定位手法が提案されている．以下，2.4 節で提案されている骨導立体音像定位手法を説明する．

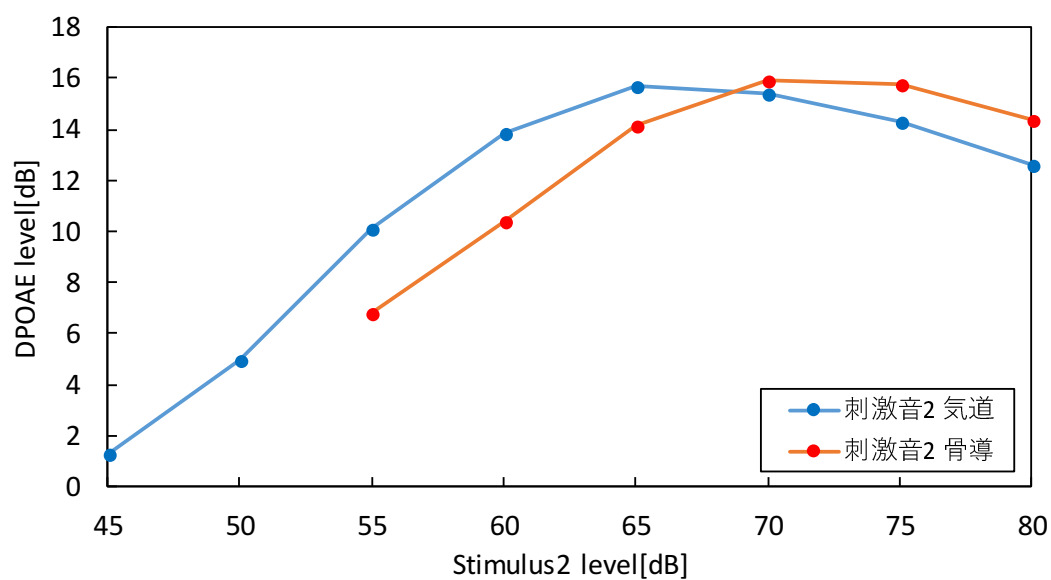


図 2.8 DPOAE の音圧変化

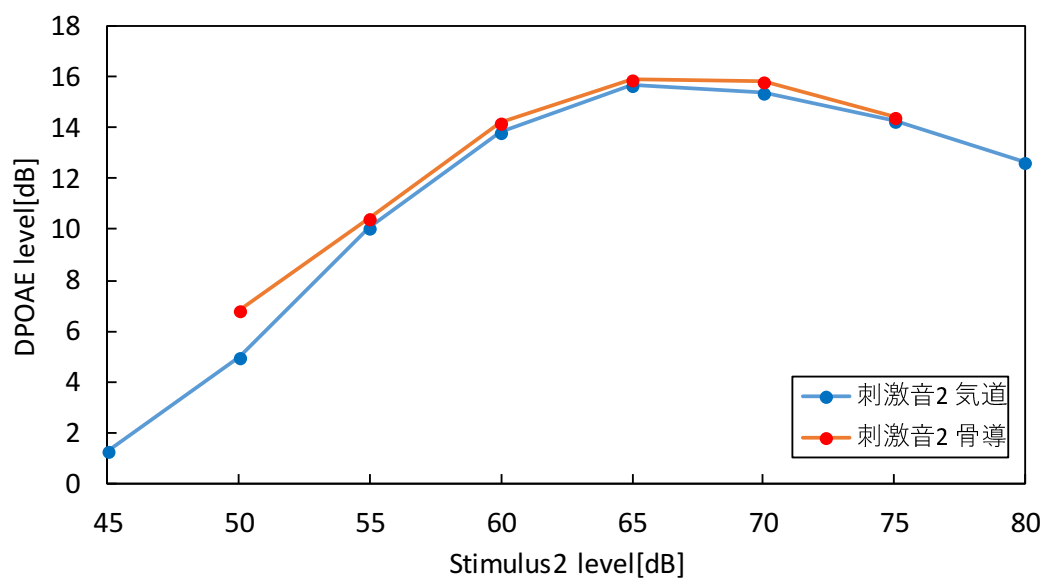


図 2.9 DPOAE の音圧変化 (シフト後)

## 2.4 骨導音による立体音像定位手法 [7]

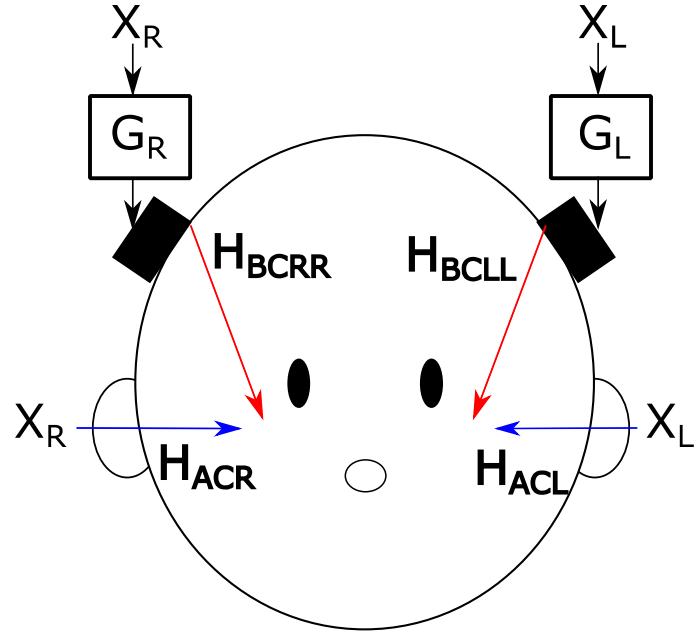


図 2.10 音の伝達経路

気道音と骨導音の伝達特性差と HRTF を組み合わせた、骨導音による立体音像定位手法を説明する．図 2.10 に音の伝達経路を示す．図 2.10 において， $H_{ACR}(z), H_{ACL}(z)$  は右（左）の外耳道入り口から右（左）蝸牛までの伝達関数， $H_{BCRR}(z), H_{BCLL}(z)$  は右（左）振動子から右（左）蝸牛までの伝達関数である

ここで， $X_R(z), X_L(z)$  は外耳道入り口において再生される信号  $x_R(t), x_L(t)$  の  $z$  変換であり， $G_R(z), G_L(z)$  は，気道音と骨導音を等価にするための前処理フィルタである．このとき，外耳・中耳を経由し蝸牛に到達する信号の  $z$  変換  $Z_{ACR}(z), Z_{ACL}(z)$  は，以下のようになる．

$$Z_{ACR}(z) = X_R(z)H_{ACR}(z) \quad (2.2)$$

$$Z_{ACL}(z) = X_L(z)H_{ACL}(z) \quad (2.3)$$

一方，左右の振動子で再生した信号が骨を経由して蝸牛に到達する信号を  $Z_{BCR}(z), Z_{BCL}(z)$  とすると，以下のようになる．

$$Z_{BCR}(z) = X_R(z)G_R(z)H_{BCRR}(z) \quad (2.4)$$

$$Z_{BCL}(z) = X_L(z)G_L(z)H_{BCLL}(z) \quad (2.5)$$

今, Rch の信号について考える. 式 (2.2) の気道音  $Z_{ACR}(z)$  と式 (2.4) の骨導音  $Z_{BCR}(z)$  を等価にする前処理フィルタ  $G_R(z)$  は, 以下のように求めることができる.

$$\begin{aligned} X_R(z)H_{ACR}(z) &= X_R(z)G_R(z)H_{BCRR}(z) \\ G_R(z) &= \frac{H_{ACR}(z)}{H_{BCRR}(z)} \end{aligned} \quad (2.6)$$

Lch についても同様に考えると以下のようなになる.

$$G_L(z) = \frac{H_{ACL}(z)}{H_{BCLL}(z)} \quad (2.7)$$

式 (2.6)-(2.7) より, 前処理フィルタ  $G_R(z), G_L(z)$  は, それぞれ骨導音に対する気道音の伝達特性差  $H_{ACR}(z)/H_{BCRR}(z)$ ,  $H_{ACL}(z)/H_{BCLL}(z)$  となること分かる.

骨伝導ヘッドホンによる音像定位精度は, 上記の前処理フィルタを用いた場合 49.2 % であり, 前処理なしの場合 32.5 % に比べ音像定位精度が向上することが示されている. したがって, 上記の手法は音像定位精度向上に効果的であるといえる. しかし, HRTF や DPOAE の測定には時間がかかり, 利用者にとって大きな負担があるという問題がある.

## 2.5 スパースモデル [18, 19]

本節では, 本論文で用いるスパースモデルと呼ばれる数学的モデルを説明する. スパースモデルとは, 膨大なデータの統計構造を少数の因子を用いて表現する手法である. ある説明因子を  $N$  次元の信号  $\mathbf{x} = (x_1, x_2, \dots, x_N)$  と表し, この信号が作用  $\mathbf{A}$  を介して  $M$  次元の観測信号  $\mathbf{y} = (y_1, y_2, \dots, y_M)$  を与えるモデルを考える. ここで, 作用  $\mathbf{A}$  が線形混合行列であるとする, 説明因子  $\mathbf{x}$  と観測信号  $\mathbf{y}$  は  $\mathbf{y} = \mathbf{Ax}$  と表せる. ただし,  $\mathbf{A}$  は  $M \times N (M < N)$  の行列である. 行列  $\mathbf{A}$  が与えられた際に, 観測信号  $\mathbf{y}$  から説明因子  $\mathbf{x}$  を推定することを考える. この推定問題は, 未知数の個数が式の個数よりも多い劣決定系である. そのため, 推定解を得るために, 制約条件を付与する. スパースモデリングでは,  $\mathbf{x}$  が 0 の要素が多い行列である. つまり, 観測信号  $\mathbf{y}$  に現れる事象は,  $\mathbf{x}$  に含まれる少数の非ゼロ要素に対応する基底から生成されることを仮定している.  $\mathbf{x}$  がほとんどゼロの要素であるという仮定より, この方程式の解を算出することができる.

基底変換により, 画像や音声スパースに表現できることがわかっている. このことから, スパース推定は画像のノイズ除去 [20] などにも用いられる.

## 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．ヒトが音を受聴する際に，音源が外耳道入り口に到達するまでの伝達関数が HRTF である．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 HRTF を組み合わせることで，受聴音に位置情報を付与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ある．この伝達特性差を用いた，骨伝音による立体音像定位手法が提案されている．しかし，HRTF と DPOAE の測定には多大な時間を要し，利用者の負担が大きいという問題がある．提案法で用いるスパースモデルとは，あるデータの特徴を小数の説明因子で表現するモデルのことである．

## 第 3 章 HRTF 合成手法

この章では、部分的に測定した HRTF から未測定方向の HRTF を合成する手法を示す。3.1 節では、従来法の主成分分析による基底展開を用いた HRTF 合成手法を示す。次に、3.2 節では、提案法であるスパース推定と誤差推定を用いた手法を示す。3.3 節では、HRTF データベースを用いて HRTF を合成し、合成精度に関して従来法と比較する。

### 3.1 主成分分析を用いた HRTF 合成 [13]

この節では、主成分分析を用いて部分的に測定した HRTF から未測定方向の HRTF を合成する文献 [13] の方法を説明する。ある被験者の HRTF は、主成分と重みの積和で合成することができる。

$i$  番目の被験者の  $j$  番目の方向の HRTF 振幅応答を要素とするベクトルを  $\mathbf{h}_{i,j}$  とし、測定方向数を  $N_d$  とすると  $\mathbf{h}_i$  は

$$\mathbf{h}_i = [\mathbf{h}_{i1} \ \mathbf{h}_{i2} \ \cdots \ \mathbf{h}_{iN_d}]^T \quad (3.1)$$

で表される。平均 HRTF 振幅応答ベクトル  $\bar{\mathbf{h}}$  を

$$\bar{\mathbf{h}} = \frac{1}{N_s} \sum_{i=1}^{N_s} \mathbf{h}_i \quad (3.2)$$

とすると、変量ベクトル  $\hat{\mathbf{h}}_i$  は次式で表せる。

$$\hat{\mathbf{h}}_i = \mathbf{h}_i - \bar{\mathbf{h}} \quad (3.3)$$

$N_s$  は分析に用いた変量ベクトル  $\hat{\mathbf{h}}_i$  の数であり、主成分分析に用いた被験者数である。変量ベクトル  $\hat{\mathbf{h}}_i$  のすべての組み合わせに対して、共分散を計算し、分散・共分散行列を算出する。分散・共分散行列の固有ベクトルが主成分行列  $\mathbf{C}$  であり、固有値が寄与率となる。この主成分行列  $\mathbf{C}$  を用いて、 $k$  番目の被験者の全方向の HRTF 振幅応答  $\mathbf{g}'_k$  を次式により合成する。

$$\mathbf{g}'_k = \mathbf{C}\boldsymbol{\omega}_k + \bar{\mathbf{h}} \quad (3.4)$$

ここで、 $\boldsymbol{\omega}_k$  は各主成分に対応する重み係数ベクトルである。 $k$  番目の被験者の HRTF 振幅応答  $\mathbf{g}_k$  の変量ベクトルを部分変量ベクトル  $\hat{\mathbf{g}}_{k_d}$  とし、次式で表す。

$$\hat{\mathbf{g}}_{k_d} = [\hat{\mathbf{g}}_{k_{d_1}} \ \hat{\mathbf{g}}_{k_{d_2}} \ \cdots \ \hat{\mathbf{g}}_{k_{d_n}}]^T \quad (3.5)$$



ここで、 $d_l$  は測定した  $n$  方向のうち  $l$  番目の方向である。

式 (3.6) で示すように、主成分行列  $\mathbf{C}$  を全方向数  $N_d$  の部分行列に分割すると、部分行列主成分行列  $\mathbf{C}_d$  は式 (3.7) のようになる。

$$\mathbf{C} = [\mathbf{C}_1 \mathbf{C}_2 \cdots \mathbf{C}_{N_d}]^T \quad (3.6)$$

$$\mathbf{C}_d = [\mathbf{C}_{d_1} \mathbf{C}_{d_2} \cdots \mathbf{C}_{d_n}]^T \quad (3.7)$$

このとき、重み係数ベクトル  $\boldsymbol{\omega}_k$  は、近似的に式 (3.8) で計算できる。

$$\boldsymbol{\omega}_k = \mathbf{C}_d \setminus \hat{\mathbf{g}}_{k_d} \quad (3.8)$$

したがって、式 (3.4) を用いて未測定方向の HRTF を合成できる。重み係数ベクトル  $\boldsymbol{\omega}_k$  を式 (3.4) に代入する際に、寄与率の大きい主成分行列  $\mathbf{C}$  を組み合わせる。

## 3.2 提案法

この節では、主成分分析とスパース推定を用いた合成手法と誤差推定を用いた合成手法を説明する。

### 3.2.1 提案法 1-1 (PCA+スパース)

従来の合成法は、寄与率の大きい主成分の組み合わせにより HRTF を合成する。寄与率の大きい主成分は、全データ中の分散が大きいものである。ここで、次元は被験者数であるため、寄与率の大きい主成分はより多くの被験者の共通の特徴を含んでおり、個人差が表れにくいといえる。HRTF の個人の特徴を表現するためには、寄与率に依存しない主成分の選択が望ましいと考える。従来法より、被験者の HRTF は少数個の主成分の組み合わせで合成ができると仮定すると、使われていない主成分の重みは 0 である。したがって、重み係数ベクトル  $\boldsymbol{\omega}_k$  はスパースであるといえる。

下記のスパース推定を解くことで、重み係数ベクトル  $\boldsymbol{\omega}_k$  を算出する。

$$\arg \min_{\boldsymbol{\omega}_k} \|\boldsymbol{\omega}_k\|_0 \quad (3.9)$$

$$\text{subject to } \hat{\mathbf{g}}_{k_d} = \mathbf{C}_d \boldsymbol{\omega}_k \quad (3.10)$$

上記の問題を解くために、本論文では文献 [18] の OMP アルゴリズムを用いた。図 3.1-3.2 に PCA と PCA+スパースの模式図を示す。主成分行列  $\mathbf{C}$  の灰色の部分は使用されない主成分であり、オレンジ色の部分が HRTF 合成に用いる主成分である。提案法 1-1 では、寄与率に依存せずに、重み係数ベクトル  $\boldsymbol{\omega}_k$  に依存し主成分の選択を行う。

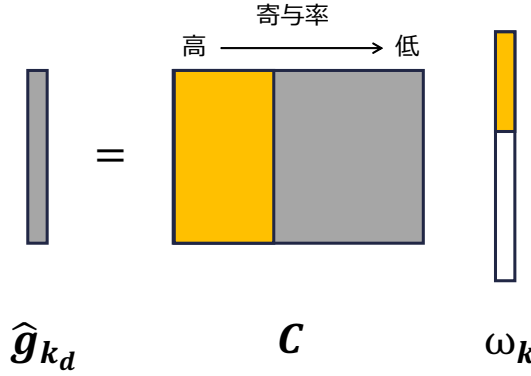


図 3.1 主成分分析模式図

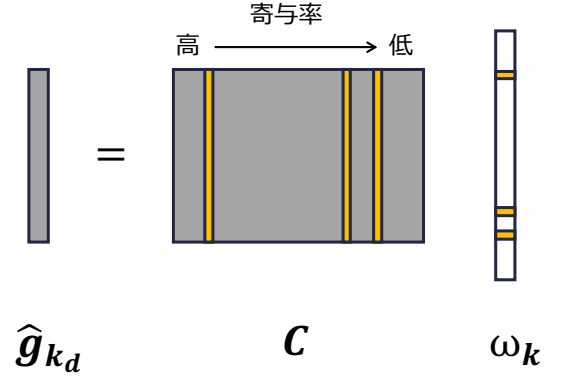


図 3.2 PCA+スパース模式図

### 3.2.2 提案法 1-2(誤差推定)

3.1 節と 3.2.1 節の手法では、実際に測定した  $n$  方向に関して被験者本人の HRTF である。しかし、合成した HRTF 振幅応答  $\mathbf{g}'_k$  は、実際に測定した  $n$  方向に関する誤差を生じる。

そこで提案法 1-2 では、主成分分析で合成した HRTF 振幅応答  $\mathbf{g}'_k$  と部分的に測定した HRTF 振幅応答  $\mathbf{g}_k$  との部分的な誤差  $\boldsymbol{\varepsilon}_k$  に着目する。

誤差  $\boldsymbol{\varepsilon}_k$  を

$$\boldsymbol{\varepsilon}_k = [\boldsymbol{\varepsilon}_{kd_1} \ \boldsymbol{\varepsilon}_{kd_2} \ \cdots \ \boldsymbol{\varepsilon}_{kd_n}]^T \quad (3.11)$$

とし、部分的に測定した  $l, l+1$  番目の方向の HRTF 振幅応答から次式で得る。

$$\boldsymbol{\varepsilon}_{kd_l} = \mathbf{g}_{kd_l} - \mathbf{g}'_{kd_l} \quad (3.12)$$

測定していない  $l'$  番目 ( $l < l' < l+1$ ) の方向の誤差  $\boldsymbol{\varepsilon}'_{kl'}$  は、次式を用いて線形補間する。

$$\boldsymbol{\varepsilon}'_{kl'} = r \times \boldsymbol{\varepsilon}_{kd_l} + (1 - r) \times \boldsymbol{\varepsilon}_{kd_{l+1}} \quad (3.13)$$

ここで、 $r$  は  $\boldsymbol{\varepsilon}_{kd_l}$  と  $\boldsymbol{\varepsilon}_{kd_{l+1}}$  の間の内分比を表すパラメータである。推定 HRTF 振幅応答  $\mathbf{g}''_k$  は、合成 HRTF 振幅応答  $\mathbf{g}'_k$  と誤差  $\boldsymbol{\varepsilon}'_k$  の和で得る。

$$\mathbf{g}''_k = \mathbf{g}'_k + \boldsymbol{\varepsilon}'_k \quad (3.14)$$

図 3.3 に従来法及び提案法の処理を示す。

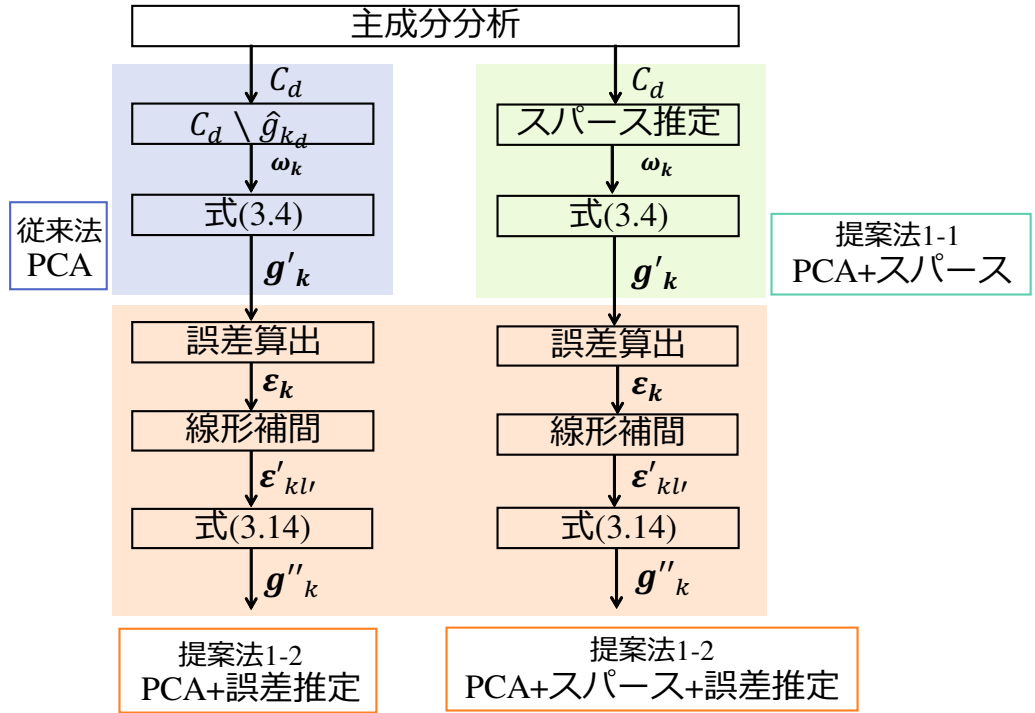


図 3.3 HRTF 合成手法

### 3.3 実験

始めに、実験条件と評価方法について説明する．HRTF は、東北大学電気通信研究所のデータベース [21, 22] を用いた．水平面内の  $0^\circ \sim 350^\circ$  まで  $10^\circ$  刻みで 36 方向，80 名分の HRTF を主成分分析の対象とした．分析に用いた 80 名以外の 20 名の水平面内の 36 方向の HRTF を評価対象とした．また，SN 比を考慮し，各方向 256 サンプルの HRTF 振幅応答のうち 17kHz 以下の 182 サンプルを分析対象とした．評価指標として次式のスペクトル歪 (SD: Spectral Distortion) の値  $S_D$  を用いる．

$$S_D = \sqrt{\frac{1}{N_f} \sum_{i=1}^{N_f} (20 \log_{10} \frac{g_k(i)}{g'_k(i)})^2} \quad (3.15)$$

ここで， $g_k(i)$  はある被験者  $k$  の測定した HRTF 振幅応答の各要素を， $g'_k(i)$  は合成した HRTF 振幅応答の各要素を， $N_f$  は HRTF 振幅応答のベクトル長であり， $N_f = 182$  である．

図 3.4-3.8 にある被験者の測定値と測定方向 4 で合成した HRTF 振幅応答を示す．図 3.9 にその時の PCA と PCA+スパースの重み係数ベクトル  $\omega_k$  を示す．図 3.6-3.6 より提案手法を用いた場合も未測定方向の HRTF 合成が可能であるといえる．また，図 3.9 より従来

法と提案法で重み係数ベクトル  $\omega_k$  は異なっていることが確認できる．PCA(主成分 1) 及び PCA+スパース (主成分 1) とそれぞれの誤差と推定誤差は付録 A.1 を参照．図 3.10 に，PCA+スパースの測定方向数，主成分数ごとの平均スペクトル歪を示す．横軸は測定方向数で，縦軸はスペクトル歪を表す．

図 3.10 より，測定方向数を削減した場合，主成分数を減少させると  $S_D$  が小さくなる．測定方向数を増やした場合，主成分数を増加させると  $S_D$  は小さくなる．図 3.11 に測定方向数 12 以下の場合の手法，主成分数ごとの平均スペクトル歪を示す．図 3.11 より， $S_D$  が最も小さくなる時の主成分数は，PCA では 15 であり，PCA+スパースでは 10 である．また，誤差推定を用いた手法では，主成分数 1 のときに  $S_D$  が最も小さい．各手法において  $S_D$  が最も小さくなる時の主成分数を用いた場合の測定方向数ごとのスペクトル歪を図 3.12 に示す．図 3.12 より，PCA と PCA+スパースは  $S_D$  の値がほぼ同じであり，未測定方向の HRTF の特性を推定するのに合成精度が同等であるといえる．また，誤差推定を用いた 2 つの手法は，誤差推定なしの手法に比べ  $S_D$  が小さい．したがって，誤差推定は未測定方向の HRTF の合成精度の向上に効果があるといえる．

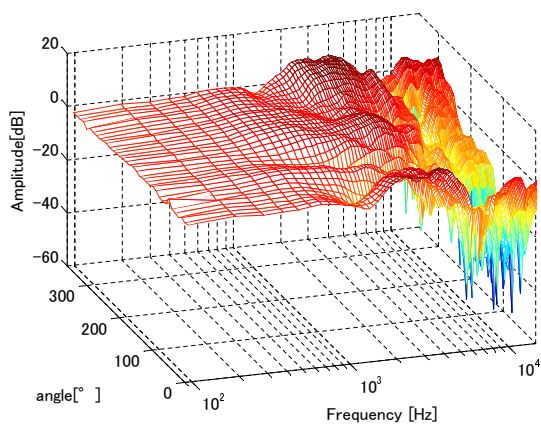


図 3.4 HRTF 測定値

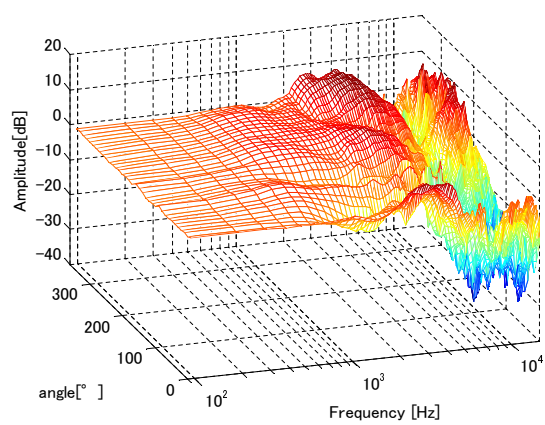
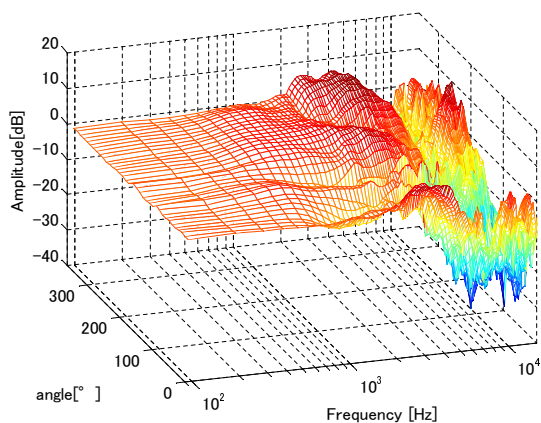
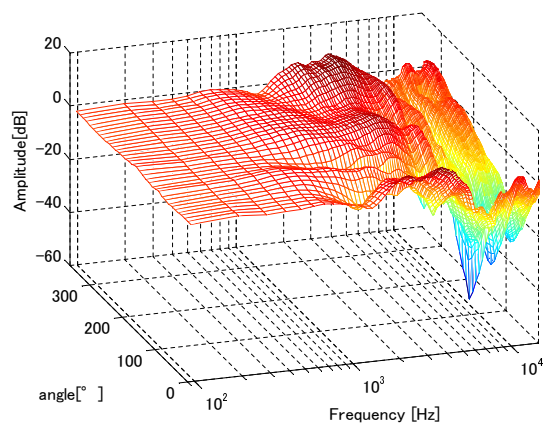
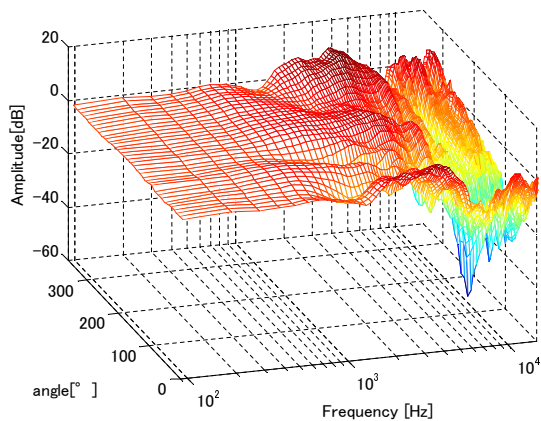
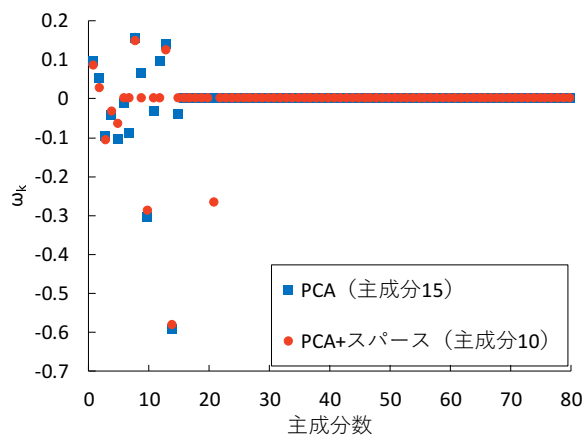


図 3.5 PCA (主成分 15)

図 3.6 PCA+スパース  
(主成分 10)図 3.7 PCA+誤差推定  
(主成分 1)図 3.8 PCA+スパース+誤差推定  
(主成分 1)図 3.9 重み係数ベクトル  $\omega_k$

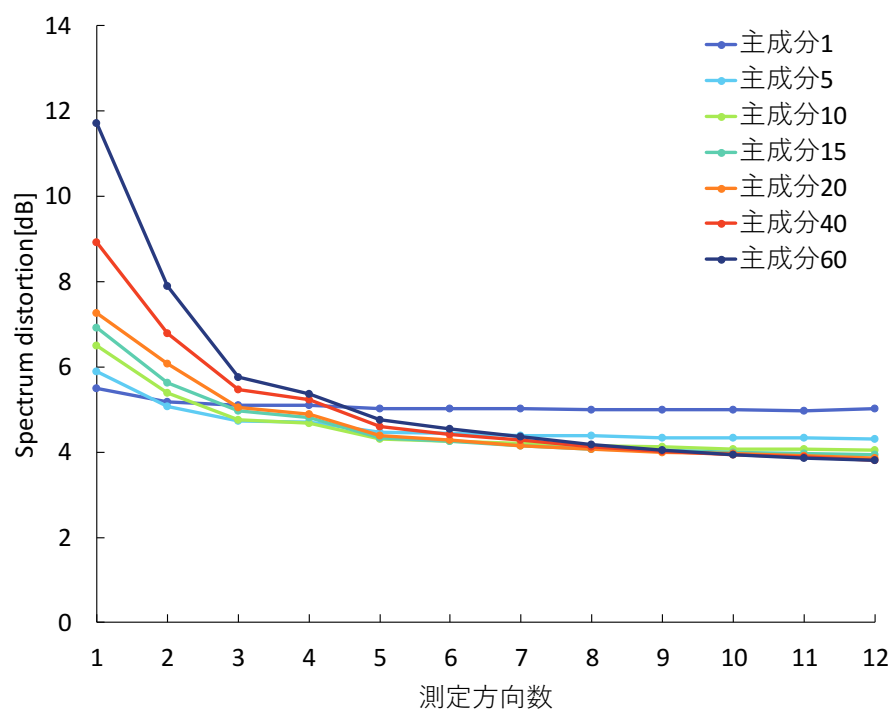


図 3.10 平均スペクトル歪 (PCA+スパース)

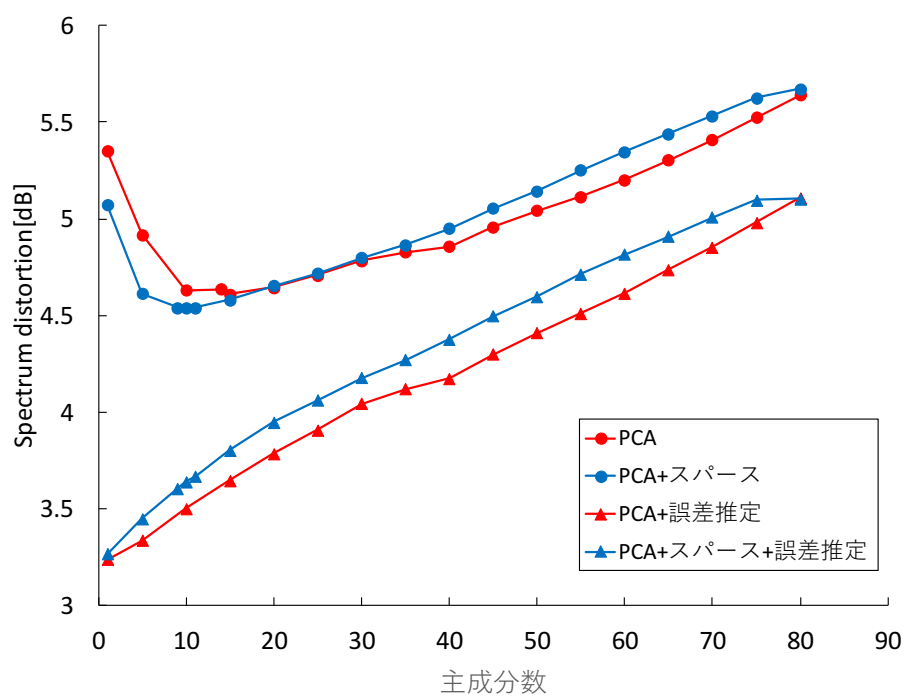


図 3.11 主成分数ごとの平均スペクトル歪

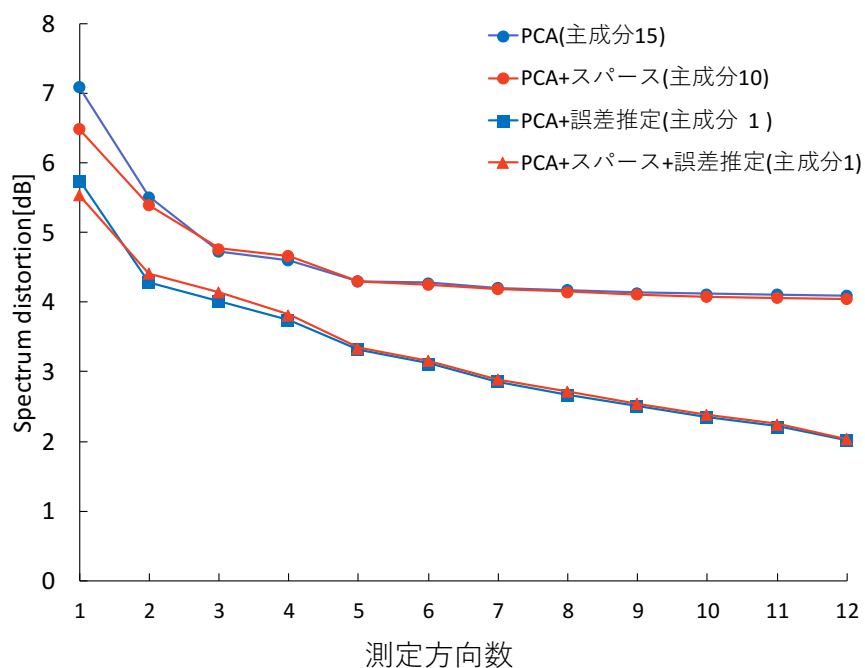


図 3.12 測定方向数ごとの平均スペクトル歪

### 3.4 まとめ

この章では、未測定方向の HRTF を合成する手法を示した。3.1 節では、主成分分析 (PCA) を用いた従来法を示した。3.2 節では、スパースモデルを用いた提案法と誤差推定を用いた提案法を示した。3.3 節では、HRTF データベースのデータを用いて、従来法と提案法の合成精度の比較を行った。誤差推定を用いた手法は、誤差推定なしの場合に比べ合成精度を示すスペクトル歪が小さい。したがって、誤差推定用いた手法は、少ない測定方向から未測定方向の HRTF を合成するのに有効であるといえる。

## 第 4 章 DPOAE 抽出手法

この章では、DPOAE 抽出手法について述べる。4.1 節では、DPOAE の周波数解析で一般的に用いられている従来法について述べる。4.2 節では、スパース推定を用いた提案法について述べる。4.3 節では、DPOAE 疑似データ及び実測データを用いた、提案法と従来法の抽出精度比較を示す。

### 4.1 DPOAE 抽出手法 (DFT)

DFT (Discrete Fourier Transform) は、一般的な周波数解析に用いられる手法である。DPOAE 録音信号をフレームごとに分割し、周波数解析をする。ここで、DPOAE 録音信号をフレーム長  $N$  で  $k$  個に分割した時間データを  $\mathbf{s}_k = [s_k(1) s_k(2) \cdots s_k(N)]^T$  とすると、DFT は次式で表せる。

$$\mathbf{x}_k = \mathbf{F}(\mathbf{w} \odot \mathbf{s}_k), \quad (4.1)$$

ここで、 $\mathbf{x}_k = [x_k(1) x_k(2) \cdots x_k(N)]^T$  はフーリエ係数で、 $\odot$  はアダマール積である。また、 $\mathbf{F}$  は  $N \times N$  の離散フーリエ変換行列であり、 $F_{m,n} = \exp(-2\pi jmn/N)$  である。 $\mathbf{w}$  は  $N \times 1$  の窓関数である。フレーム毎に算出したフーリエ係数  $\mathbf{x}_k$  を平均し、DPOAE 信号の周波数スペクトルを算出する。

$$\mathbf{x} = \sum_{k=1}^K |\mathbf{x}_k| \quad (4.2)$$

DFT では、周波数分解能がフレーム長  $N$  に依存する。そのため、フレーム長を短くした録音信号では周波数分解能が粗くなり、DPOAE 周波数が抽出できないといった問題が生じる。

### 4.2 提案法 2-1(スパース推定)

文献 [23] では、フーリエ係数  $\mathbf{x}_k$  がスパースである場合、時間データの外挿が可能となる手法が提案されている。図 4.1 に DPOAE 信号の周波数スペクトルを示す。DPOAE 信号の周波数スペクトルは、刺激音周波数 ( $f_1, f_2$ ) と DPOAE 周波数 ( $f_{dp} = 2f_1 - f_2$ ) 以外はノイズであり、フーリエ係数はスパースであるといえる。したがって、文献 [23] をもとに



短い録音信号から高分解能の周波数スペクトルを算出することができる．以下，文献 [23] の手法を説明する．あるフレーム  $k$  の時間データ  $\mathbf{s}_k$  に対して，外挿して得られるフレーム長  $L(> N)$  のデータを  $\hat{\mathbf{s}}_k = [\hat{s}_k(1) \hat{s}_k(2) \cdots \hat{s}_k(L)]^T$  とすると，DFT は次式で表せる．

$$\hat{\mathbf{x}}_k = \hat{\mathbf{F}}(\hat{\mathbf{w}} \odot \hat{\mathbf{s}}_k), \quad (4.3)$$

$\hat{\mathbf{x}}_k = [\hat{x}_k(1) \hat{x}_k(2) \cdots \hat{x}_k(L)]^T$  で， $\hat{\mathbf{F}}$  は  $L \times L$  の離散フーリエ変換行列を， $\hat{\mathbf{w}}$  は  $L \times 1$  の窓関数を表す．

ここで， $N \times L$  の選択行列  $\mathbf{M}$  を用いる． $\mathbf{M}$  は  $L \times L$  の単位行列  $I_{L \times L}$  の最初の  $Q/2$  行と最後の  $Q/2$  行を切り捨てたものである ( $Q = L - N$ )． $\mathbf{M}$  を用いると， $\hat{\mathbf{s}}_k$  と  $\mathbf{s}_k$  は次式で表せる．

$$\mathbf{M}\hat{\mathbf{s}}_k = \mathbf{s}_k \quad (4.4)$$

窓関数を組み入れると

$$\mathbf{M}(\hat{\mathbf{w}} \odot \hat{\mathbf{s}}_k) = (\mathbf{M}\hat{\mathbf{w}}) \odot \mathbf{s}_k \quad (4.5)$$

となる．ここで， $\mathbf{G}$  は  $L \times L$  の逆離散フーリエ変換行列とすると， $\hat{\mathbf{x}}_k$  の逆フーリエ変換は

$$\mathbf{G}\hat{\mathbf{x}}_k = \hat{\mathbf{w}} \odot \hat{\mathbf{s}}_k \quad (4.6)$$

と表せる．したがって，式 (4.5) は次式で表せる．

$$\mathbf{M}\mathbf{G}\hat{\mathbf{x}}_k = (\mathbf{M}\hat{\mathbf{w}}) \odot \mathbf{s}_k \quad (4.7)$$

このとき，それぞれのフレームに対して下記の最適化問題を解くことで，短い録音信号  $\mathbf{s}_k$  から高分解能の周波数スペクトル  $\hat{\mathbf{x}}_k$  を得ることができる．

$$\arg \min_{\hat{\mathbf{x}}_k} \|\hat{\mathbf{x}}_k\|_1 \quad (4.8)$$

$$\text{subject to } \|\mathbf{M}\mathbf{G}\hat{\mathbf{x}}_k - (\mathbf{M}\hat{\mathbf{w}}) \odot \mathbf{s}_k\|_2 \leq \varepsilon \quad (4.9)$$

ここで， $\varepsilon$  は許容誤差， $\|\cdot\|_1$  は L1 ノルムである．上記の問題を解くために，本論文では文献 [24] の SPGL1 アルゴリズムを用いた．

フレーム毎に算出したフーリエ係数  $\hat{\mathbf{x}}_k$  は，DPOAE 周波数にピークが確認されないフレームが存在する場合がある．フレーム毎のフーリエ係数を平均する際に，このようなフレームを除去することでより正確な DPOAE 音圧を抽出する．DPOAE 周波数にピークが確認されるフレーム，すなわち DPOAE 周波数に極大値をとるフレームのみを平均し，DPOAE 音圧を抽出する．

図 4.2-4.4 に DFT とスパース推定による DPOAE 抽出手法を示す．

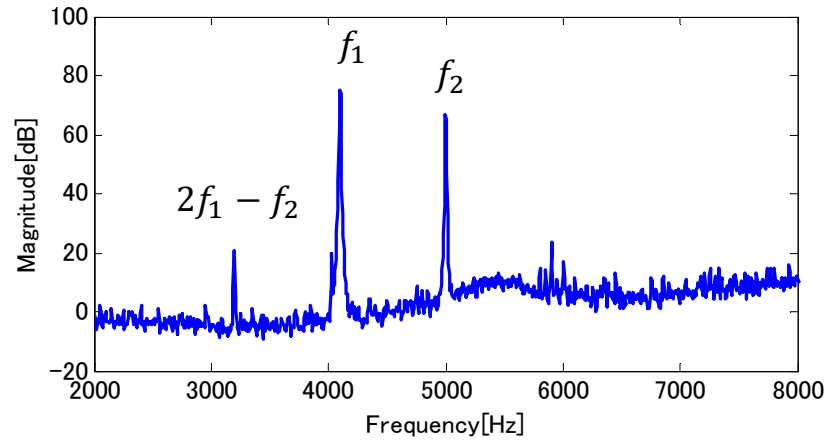


図 4.1 DPOAE スペクトル

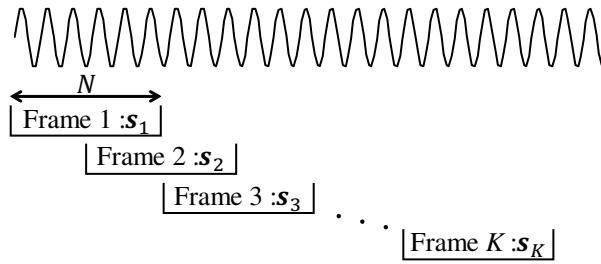


図 4.2 録音信号のフレーム分割

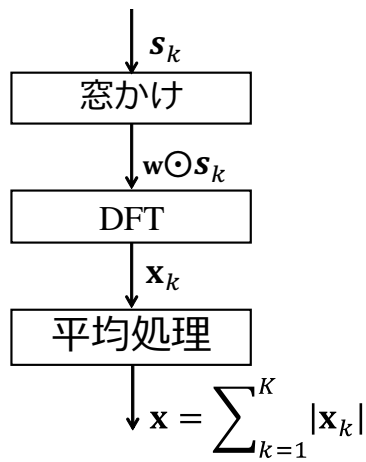


図 4.3 従来法 (DFT)

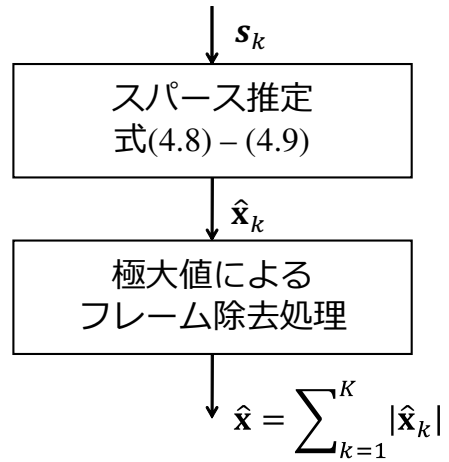


図 4.4 提案法 (スパース推定)

表 4.1 解析条件 (DPOAE 疑似データ)

サンプリング周波数	48000[Hz]
刺激音 1 音圧 $L_1$	75[dB]
刺激音 2 音圧 $L_2$	75[dB]
DPOAE 音圧 $L_{dp}$	15[dB]
$N$	1024
$L$	8192
$\varepsilon$	$10^{-4}$
$K$	10
窓関数	hann

### 4.3 実験

この節では、従来法と提案法を用いて DPOAE 抽出を行う。DPOAE 疑似データ及び実測データを用いて提案法と従来法の抽出精度を比較する。

#### 4.3.1 DPOAE 疑似データ

提案法と従来法の抽出精度を比較するために、DPOAE の疑似データを用いる。疑似データは刺激音 1,2 と DPOAE で構成される。ここで、2つの刺激音の周波数をそれぞれ  $f_1, f_2 (= 1.22 \times f_1)$  とし、DPOAE 周波数を  $f_{dp} (= 2f_1 - f_2)$  とする。ここでは、 $N = 1024$  の時間データを提案法により外挿することで、 $L = 8192$  に高分解能化することを考える。表 4.1 に解析条件を示す。図 4.5-4.7 に  $f_2$  を変化させた場合の疑似データの周波数スペクトルを示す。比較のため、DFT( $N = 1024$ ) の結果も示す。DFT による解析では提案法の周波数点  $L$  である 8192 点にそろえるために、ゼロパディングを行った。また、刺激音 1 が 75[dB] になるように正規化した。図 4.5-4.6 において、従来法では見られない DPOAE 周波数のピークが提案法により抽出できることを確認した。図 4.7 より、 $f_2$  が 3000[Hz] のときは従来法の DFT( $N = 1024$ ) で解析した場合でも DPOAE 周波数にピークが確認できる。 $f_2 \geq 3000$ [Hz] については付録 A.2 を参照。

そこで次に、解析時のフレーム長である、 $N$  と  $L$  を変更し DPOAE 音圧の調査を行った。 $N$  と  $L$  を変更したときのフレーム除去処理なしとありの場合の結果をそれぞれ表 4.2 と表 4.3 に示す。疑似データの DPOAE 音圧は 15[dB] であり、表中の値は提案法によって抽出

した DPOAE 音圧と真値との絶対値誤差 [dB] であり，赤色は，各  $f_2$  において最も誤差が小さい値である．表の黒色は従来法の  $\text{DFT}(N)$  の場合も DPOAE 周波数のピークが検出可能な条件である．また，表 4.3 の黄色はフレーム除去処理により，誤差値が小さくなった条件であり，－ はフレーム除去により全フレームが削除されたものである．表 4.2-4.3 より， $f_2$  によって正確な DPOAE 音圧を抽出する際の条件が異なることがわかる． $f_2$  を大きくすると，フレーム長  $N$  は少数で解析することができる．DPOAE 疑似データは  $f_2/f_1 = 1.22$  で， $f_{dp} = 2f_1 - f_2$  であり， $f_2$  が小さいときは  $f_1$  と DPOAE 周波数  $f_{dp}$  の間隔が狭い．そのため， $f_2$  によって正確な DPOAE 音圧を抽出するフレーム長の条件が異なると考えられる．疑似データを用いた調査において，フレーム除去処理の効果的な条件の規則性等を確認できなかった．

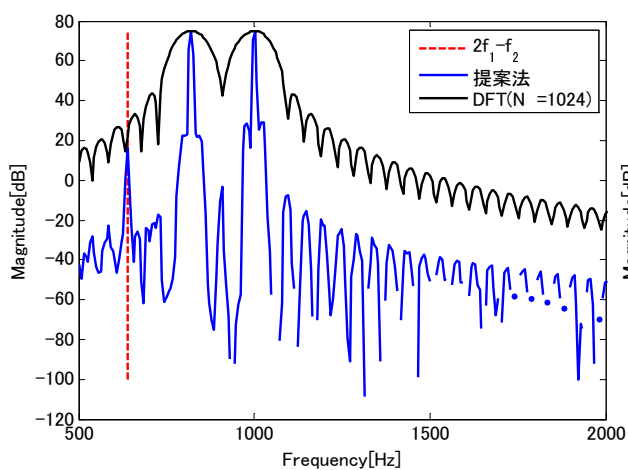


図 4.5  $f_2 = 1000Hz$

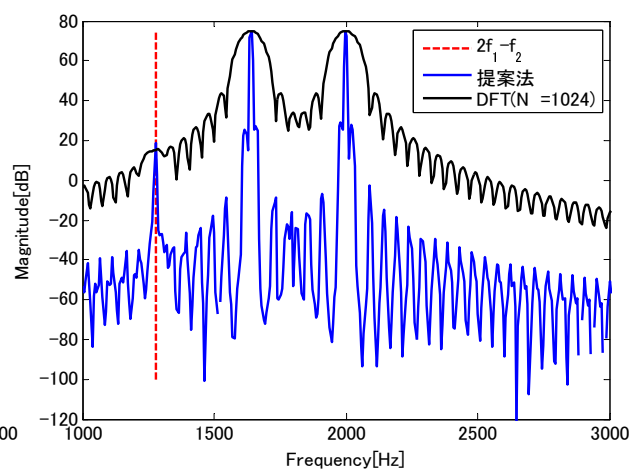


図 4.6  $f_2 = 2000Hz$

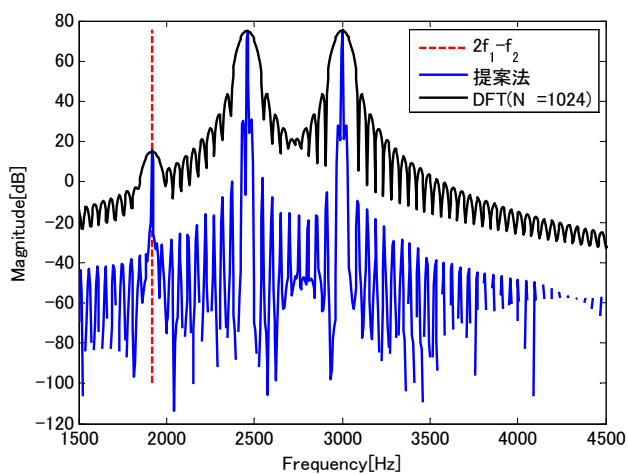


図 4.7  $f_2 = 3000Hz$

表 4.2 DPOAE 音圧誤差値 (フレーム除去処理なし)

N	1024		512			256				128				
	4096	2048	4096	2048	1024	4096	2048	1024	512	4096	2048	1024	512	256
L														
1000	8.4	2.8	2.8	0.3	10.5	9.4	1.3	2.3	31.7	31.1	26.6	41.8	46.6	46.9
2000	1.1	2.3	0.3	26.8	2.7	52.9	8.2	0.8	10.9	21.7	2.9	10.4	0.6	31.8
3000			2.6	14.2	0.3	43.7	12.4	4.6	2.6	14.7	1.5	32.7	9.5	15.2
4000			1	3.8	0.5	54.1	4.4	15.3	2.5	48.2	25.4	11.5	0.9	11.6
5000						22.9	5.4	1.4	1	0.8	1.1	0.9	42	1.6
6000						7.4	6.4	8.2	0.1	-	28.1	4.5	2.2	1.4
7000						2.2	3.8	1.3	1	2.5	0	0.5	6.5	3.9
8000						1.8	2.4	1.9	1.5	51.5	0.6	5.5	21.7	2.5
9000										5.2	5.2	0.9	10.9	4.1
10000										22.9	2.1	3.7	2	0.3

表 4.3 DPOAE 音圧誤差値 (フレーム除去処理あり)

N	1024		512			256				128				
	4096	2048	4096	2048	1024	4096	2048	1024	512	4096	2048	1024	512	256
L														
1000	1.9	4.2	1.3	5.9	13.4	-	4	-	-	-	-	-	-	-
2000	0.3	2.4	0.6	-	3.7	-	5	8.7	14.5	-	5.1	-	-	-
3000			1.1	-	3.5	-	-	3.8	5.1	-	4.1	-	9.1	-
4000			1	0.4	2.3	-	3.3	1.9	4.5	-	-	1	6.6	12.6
5000						-	4	2.4	4.4	1.3	4.7	0.5	-	4.7
6000						2	7.7	0.4	2.9	-	-	1.2	2.8	3.2
7000						3.8	3.8	2.6	0.1	4	4.2	1.6	3.2	7.9
8000						2.2	2.3	0.2	1.8	-	2.4	0.3	0.1	3.7
9000										3.3	0.5	3.8	0.7	6.1
10000										-	2.5	0.2	3	4.3

### 4.3.2 DPOAE 実測データ

表 4.4 に解析に用いる DPOAE 実測データの測定条件を示し、図 4.8 に測定の簡略図を示す。被験者が静止した状態で刺激音 1,2 を気道音で与え、各被験者の測定信号を録音した。解析条件は、フレーム  $N, L$  は疑似データの結果より表 4.5 の値である。その他の解析条件は、表 4.1 と同じ値を用いた。ここで、 $\text{DFT}(N = 8192)$  を 700 回加算平均し対ノイズ比が向上している DPOAE 音圧を真値とする。図 4.9-4.12 に被験者 4 名分の解析結果を示す。横軸は刺激音周波数  $f_2$  で、縦軸は提案法による DPOAE 音圧と真値との絶対値誤差である。また、図中の赤線はフレーム除去ありの場合の DPOAE 音圧の誤差値であり、データがプロットされていないデータは全フレームが除去されたことを示す。図 4.13 に subjectA,  $f_2 = 2000[\text{Hz}]$  の周波数スペクトルを示す。提案法と同じフレーム長の  $N = 512$  で周波数解析を行った従来法は、フレーム除去処理なしに比べて DPOAE 音圧の真値に近い値をとる。しかし、DPOAE 周波数部分にピークが見られず、DPOAE 発生の有無を確認することができない。一方、提案法のフレーム除去処理ありでは、DPOAE 周波数部分に明かなピークを確認できる。また、真値に近い値が得られている。

subjectA と B では、ほとんどの  $f_2$  でフレーム除去処理ありは処理なしに比べて誤差が小さくなり、フレーム除去処理による効果が確認できる。しかし、subjectC と D では一部の  $f_2$  でフレーム除去処理による効果は確認できるが、全フレームが除去された  $f_2$  も存在する。全フレームが除去された場合に関して、解析範囲内に DPOAE が発生しているか確認した。提案法では、録音信号をフレーム長  $N$ 、フレーム数  $k(= 10)$ 、ハーフオーバーラップで分割しており、解析範囲は  $(2/N) \times 11$  である。 $\text{DFT}(k = 1)$  を用いて、この解析範囲の周波数解析を行った。図 4.14 にフレーム除去処理で全フレームが除去された subjectD,  $f_2 = 6000[\text{Hz}]$  の周波数スペクトルを示す。図 4.14 より、提案法の解析範囲  $((2/N) \times 11 = 1408)$  に DPOAE 周波数のピークは確認されなかった。したがって、解析範囲内において DPOAE が発生していない、又は、DPOAE がノイズよりも小さいといえる。このような録音信号は、フレーム除去処理において全フレームが除去されたといえる。

表 4.4 DPOAE 測定条件

サンプリング周波数	48000[Hz]
被験者数	4 人
刺激音長	120[sec]
刺激音 2 周波数 $f_2$ 範囲	1000Hz - 10000Hz
刺激音 2 周波数 $f_2$ 間隔	1000Hz
刺激音 1 音圧 $L_1$	75dB
刺激音 2 音圧 $L_2$	65dB
刺激音出力機器	Audio-technica HEADPHONE AMPLIFIER AT-HA21
マイク	Etymotic Reseach ER-10B

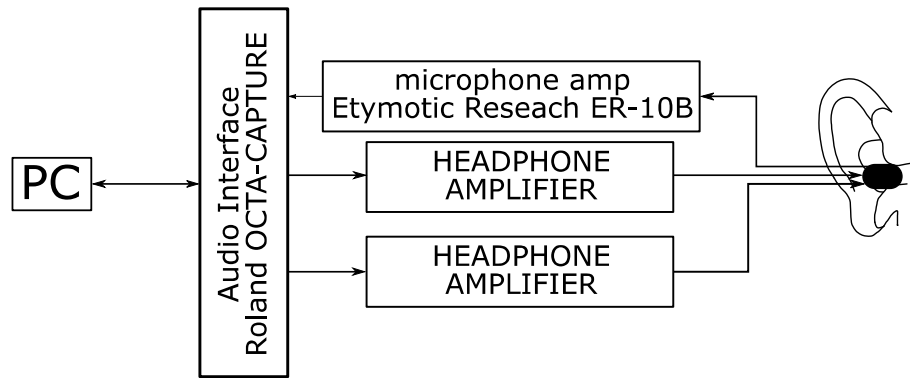


図 4.8 DPOAE 測定の簡略図

表 4.5 解析条件 (DPOAE 実測データ)

$f_2$	$N$	$L$
1000	512	4096
2000	512	4096
3000	512	4096
4000	512	2048
5000	256	1024
6000	256	1024
7000	256	512
8000	256	1024
9000	128	2048
10000	128	1024



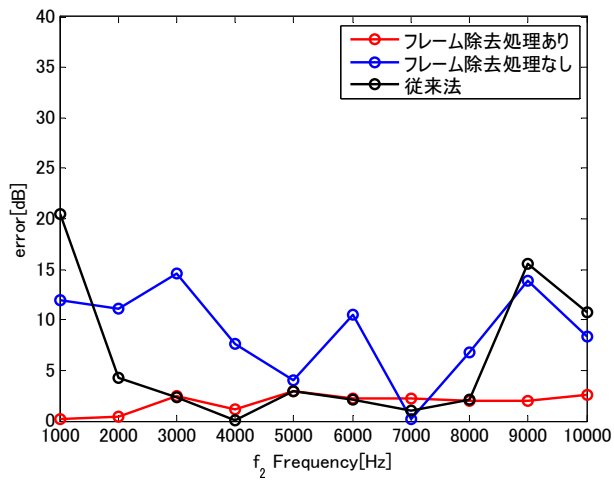


図 4.9 subjectA

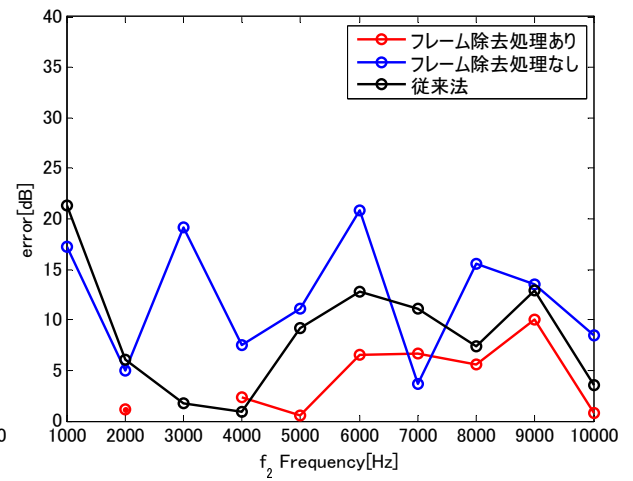


図 4.10 subjectB

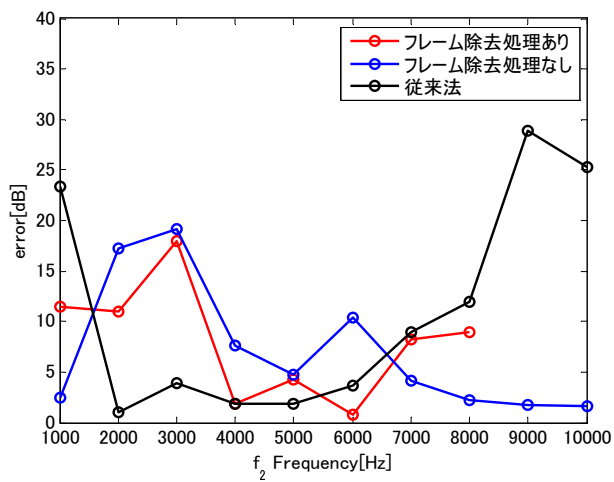


図 4.11 subjectC

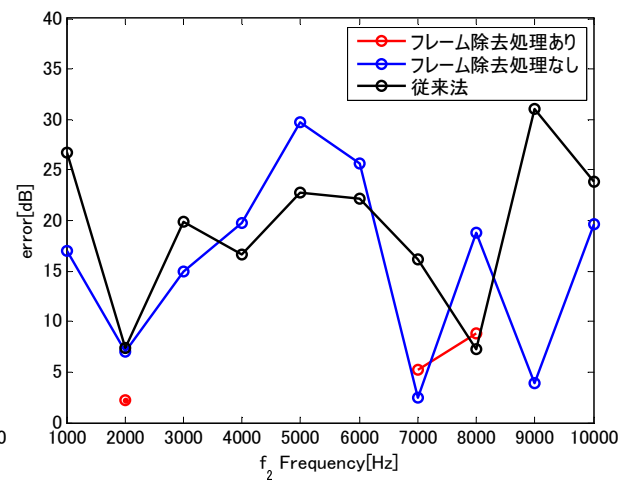
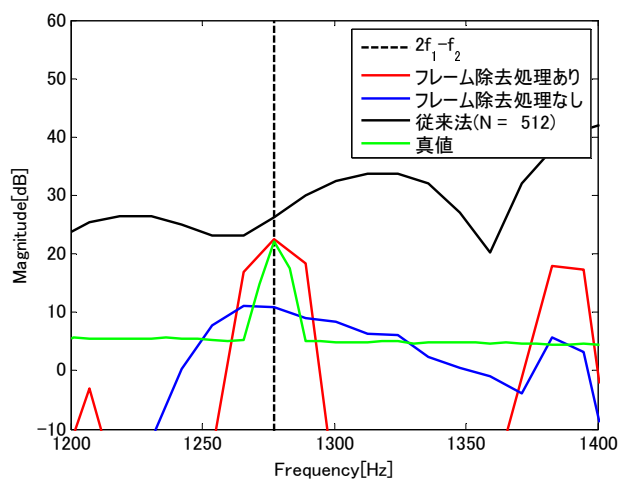
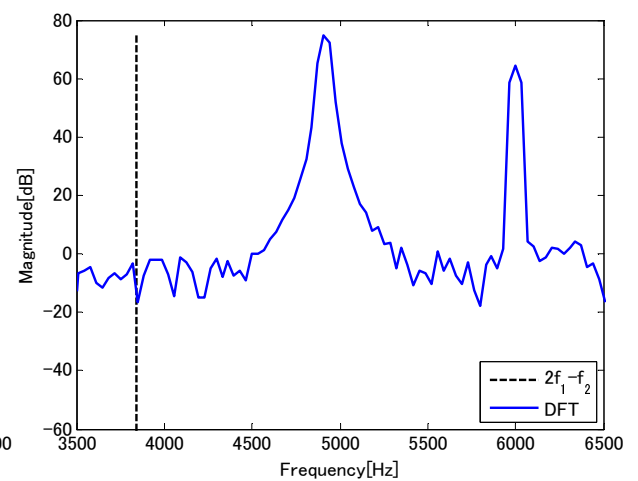


図 4.12 subjectD

図 4.13 subjectA  $f_2 = 2000$  [Hz]図 4.14 subjectB  $f_2 = 6000$  [Hz]

## 4.4 まとめ

この章では，従来法と提案法による DPOAE 抽出手法について述べた．DFT は一般的に DPOAE の解析で用いられている．しかし，周波数の分解能を向上させるためには，測定時間を長くする必要があり，被験者に多大な負荷がかかるという問題点が生じる．スパース推定を用いて短い測定時間からより正確な DPOAE 音圧を抽出するために，極大値によるフレーム除去処理を提案した．実験では，疑似データ及び実測データを用いて提案法と従来法の抽出精度を比較し，提案法の効果を検証した．提案法により，測定時間を短時間化し，従来法では抽出できなかった DPOAE を抽出することができた．また，疑似データを用いて刺激音周波数によってフレーム長を変えることにより，提案法が有効である解析条件を調査した．フレーム除去処理の効果を一部で確認することができた．フレーム除去処理にて全フレームが除去されてしまうデータは，解析範囲において DPOAE 周波数のピークを確認することができなかった．

## 第 5 章 おわりに

本論文では、測定時間の短時間化を目的とし、主成分分析による基底展開とスパースモデルを用いて部分的に測定した HRTF から未測定方向の HRTF を合成する手法を提案し、従来法よりも高い合成精度を目指した。また、スパース推定を用いて短い測定時間から伝達特性差の推定に必要な DPOAE を抽出する手法を提案し、従来法では抽出ができなかった DPOAE の抽出を目指した。

第 1 章では、骨伝導ヘッドホンによる立体音像定位について述べ、本論文の研究目的を述べた。第 2 章では、本研究の基礎となる知識・手法と提案法で用いるスパース推定を述べた。また、骨導音による立体音像定位手法の問題点として、HRTF 及び DPOAE の長時間の測定を挙げた。第 3 章では、従来法及び提案法による未測定方向の HRTF を合成する手法について述べた。また、合成精度に関して従来法と比較を行った。提案法である誤差推定を用いた手法は、誤差推定なしの場合に比ベスペクトル歪が小さく、従来法よりも合成精度が向上したことを示した。提案法により、測定時間を短時間化し従来法よりも高い合成精度を得ることができた。第 4 章では、従来法及び提案法による DPOAE 抽出手法について述べた。また、疑似データ及び実測データを用いて従来法と抽出精度を比較し、提案法の効果を検証した。周波数ごとに解析幅を変える事により提案法が効果的であることを示した。より正確な DPOAE 音圧を抽出するための極大値を用いたフレーム除去処理において、全フレームが除去されるデータは解析範囲に DPOAE 発生の有無が確認されないことを示した。提案法により、測定時間を短時間化し従来法では抽出できなかった DPOAE の抽出をすることができた。

## 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坪根正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り，そして多くのご指摘を下された，杉田研究室の同期・後輩の皆様に感謝の意を表します。最後に，学生生活を常に支え続けてくれた両親に心から感謝いたします。

平成 31 年 2 月 8 日

## 付録 A 付録

付録 A.1 に 3.3 の主成分数 1 の合成 HRTF, 誤差, 及び推定誤差を示す. 付録 A.2 に 4.3.1 節の  $f_2 \geq 3000[\text{Hz}]$  の周波数スペクトルを示す.

### A.1 付録 1

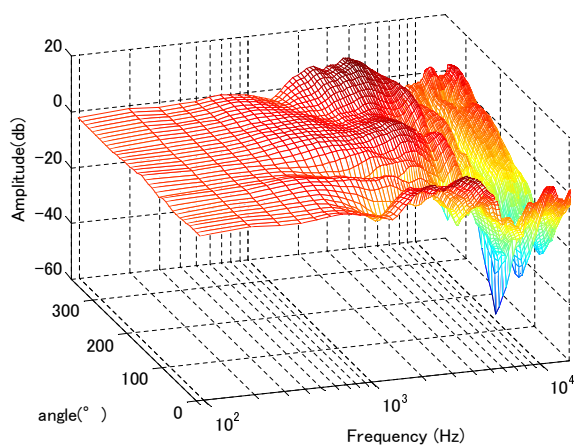


図 A.1 PCA

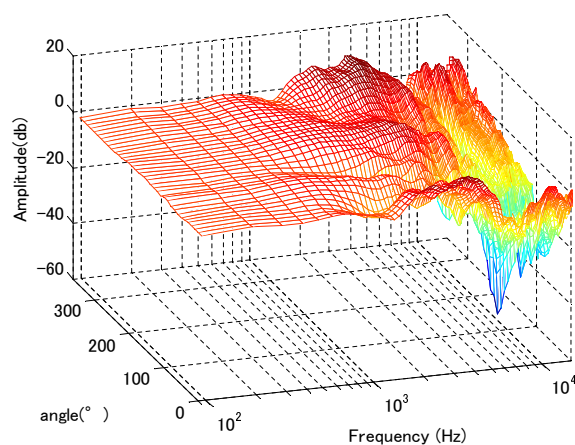


図 A.2 PCA+スパース

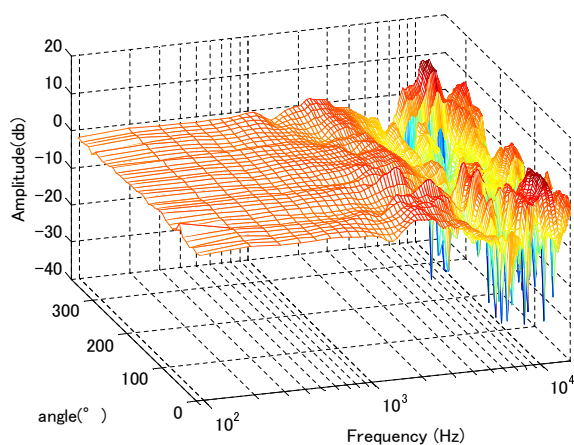


図 A.3 誤差 (PCA)

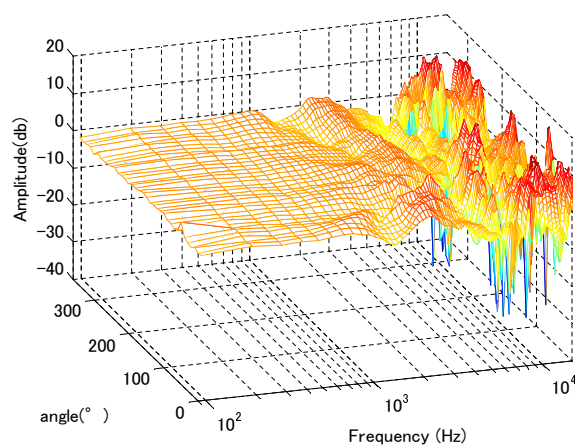


図 A.4 誤差 (PCA+スパース)

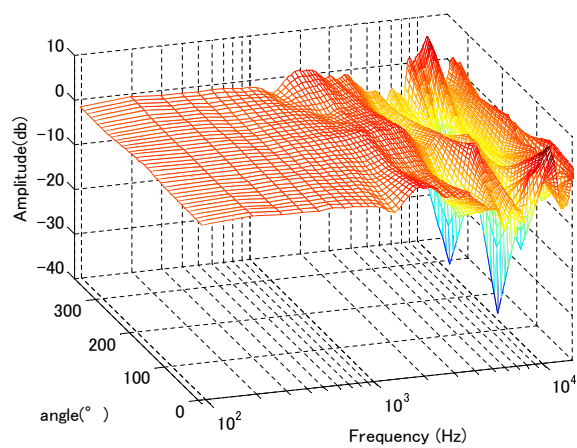


図 A.5 測定方向数4における  
推定誤差 (PCA)

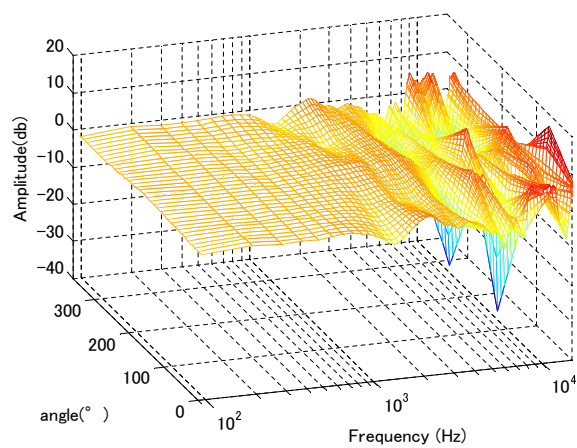


図 A.6 測定方向数4における  
推定誤差 (PCA+スパース)

## A.2 付録2

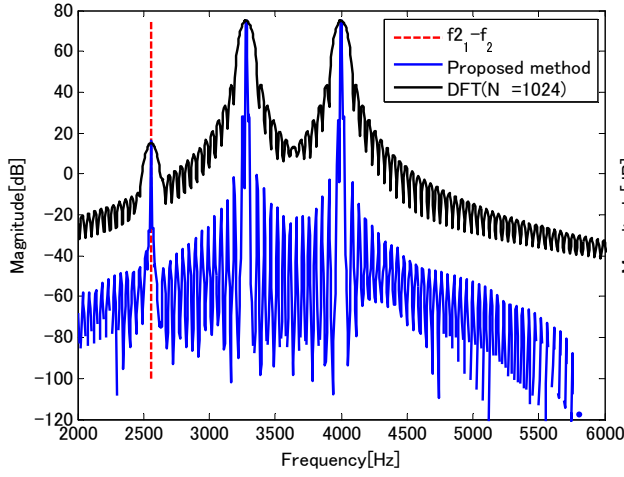


図 A.7  $f_2 = 4000Hz$

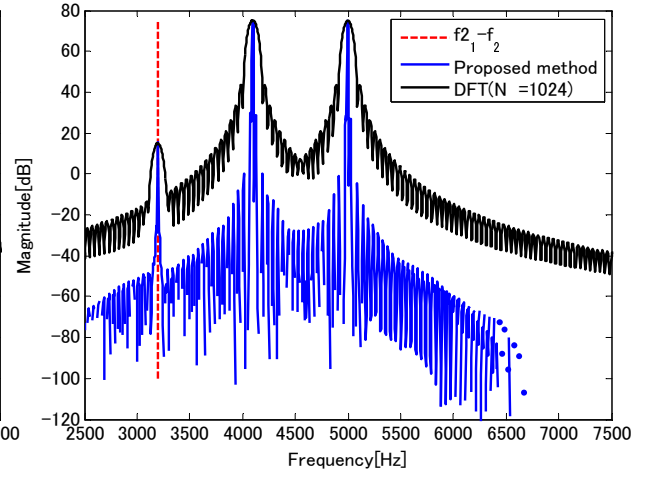


図 A.8  $f_2 = 5000Hz$

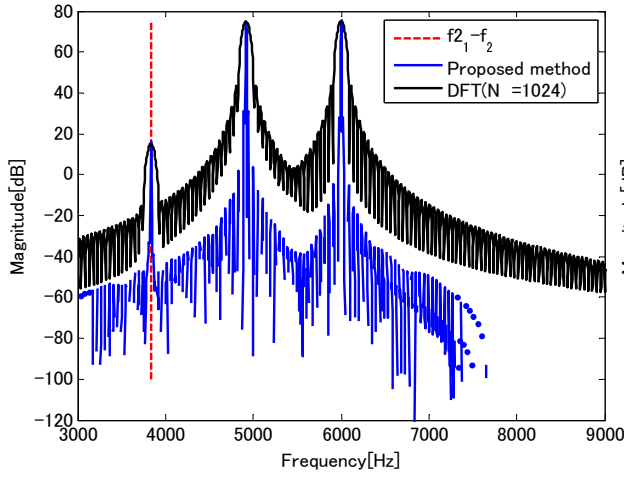


図 A.9  $f_2 = 6000Hz$

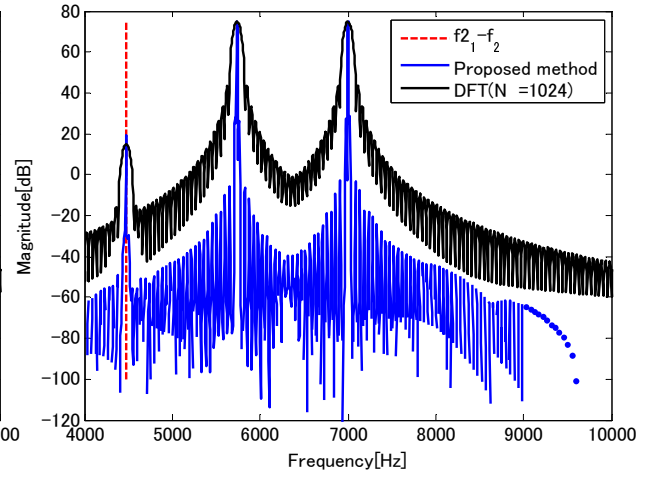


図 A.10  $f_2 = 7000Hz$

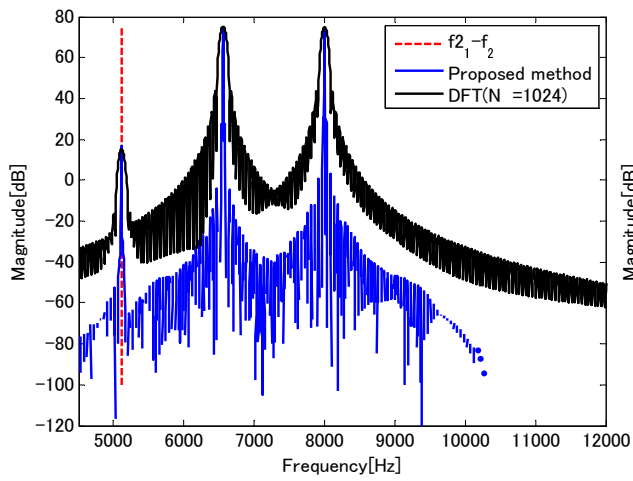


図 A.11  $f_2 = 8000Hz$

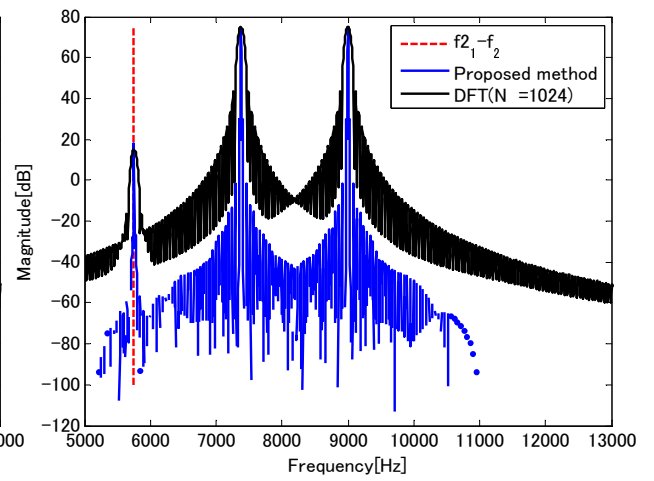


図 A.12  $f_2 = 9000Hz$

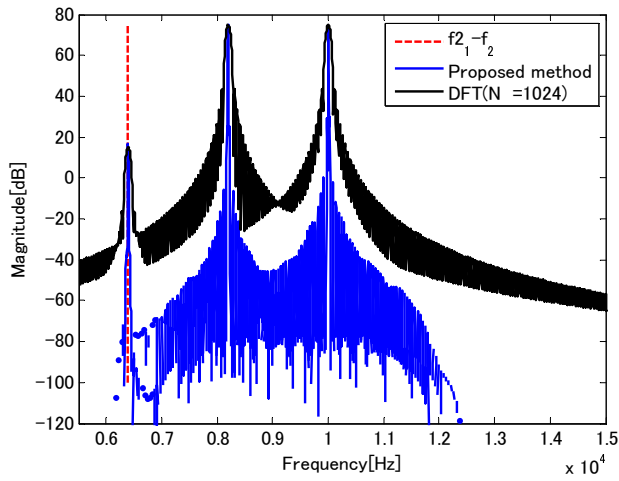


図 A.13  $f_2 = 10000Hz$



## 参考文献

- [1] M.Morimoto, Y.Ando, “On the simulation of sound localization,” Journal of the Acoustical Society of Japan, Vol.1(3), pp.167-174, May, 1980.
- [2] C.Yoshiki, M.Iwaki, “A compensation method of sound source direction perceived through bone-conduction headphones by emphasizing interaural level difference,” Consumer Electronics(GCCE), 2015 IEEE 4th Global Conference on. IEEE, pp.395-398, Feb, 2015.
- [3] Irwansyah, U.Tsuyoshi, “Application of active control technique on a bone conduction headphone for estimating a cross-talk compensation filter,” Region 10 Conference, TENCON 2017-2017 IEEE, IEEE, pp.3099-3104, Dec, 2017.
- [4] A.Barde, W.S.Helton, G.Lee, M.Billinghurst, “Binaural spatialization over a bone conduction headset: Minimum discernable angular difference,” In Audio Engineering Society Convention 140, Audio Engineering Society, May, 2017.
- [5] 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, Oct, 2011.
- [6] 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, Feb, 2011.
- [7] 多田紘大, 杉田泰則, “気道・骨導伝達特性の差を利用した骨伝導ヘッドホンの音像定位精度向上に関する-検討,” 電気学会研究会資料, CT 2017.19, pp.15-20, 2017.
- [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] 飯田一博, “頭部伝達関数の基礎と 3 次元音響システムへの応用,” 日本音響学会, コロナ社, 東京, 2017.

- [10] M.Morimoto, Y.Masayuki, A.Yoichi, “On the simulation of sound localization,” *Journal of the Acoustical Society of Japan* (e) 1.3, pp.167-174,1980.
- [11] 西野隆典, 梶田将司, 武田一哉, 板倉文忠, “水平面上の頭部伝達関数の補間,” *日本音響学会誌*, 55(2), pp.91-99, 1999.
- [12] 西野隆典, 梶田将司, 武田一哉, 板倉文忠 “水平方向及び仰角方向に関する頭部伝達関数の補間,” *日本音響学会誌*, 57(11), pp.685-692, 2001.
- [13] 松井健太郎, 安藤影男, “主成分分析に基づく個人性頭部伝達関数の推定,” *NHK 技研 R&D*, No.126.2011.3, pp.49-57, 2011.
- [14] 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,” *Annals of Otology, Rhinology & Laryngology*, 99, pp.3-14, 1990.
- [15] 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.
- [16] T.Chou, Y.Sugita, “Non-linear Features Extraction of Biometric Authentication Using System Identification,” *Proc. of The 2018 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing*, pp.513-516, 2018.
- [17] D.W.Purcell, H.Kunov, W.Cleghorn, “Estimating bone conduction transfer functions using otoacoustic emissions,” *The Journal of the Acoustical Society of America*, Vol.114(2), pp.907-918, 2003.
- [18] 庄野逸, “スパースモデリングの歴史と基本技術 (特集 スパースモデリングの発展: 原理から応用まで)–(全体概要と基本理論),” *電子情報通信学会誌, The journal of the Institute of Electronics, Information and Communication Engineers*, vol.99, No.5 pp.376-380, 2016.
- [19] M.Elad, 玉木徹 訳, “スパースモデリング,” 共立出版, 東京, 2016.
- [20] J.Mairal, E.Michael, S.Guillermo “Sparse representation for color image restoration,” *IEEE Transactions on image processing*, vol.17.1, pp.53-69 , Jan, 2008

- [21] 東北大学電気通信大学先端音響情報システム研究室, : The RIEC HRTF Dataset, <http://www.riec.tohoku.ac.jp/pub/hrtf/index.html>
- [22] K.Watanabe, Y.Iwaya, Y.Suzuki, S.Takane, S.Sato, “Dataset of head-related transfer functions measured with a circular loudspeaker array,” *Acoustical science and technology*, 35(3), pp.159-165, 2014.
- [23] J.Akhtar, K.E.Olsen. “Sparse Frequency Extrapolation of Spectrograms,” *Proc. of 25th EUSIPCO*, pp.1175-1179, 2017.
- [24] E.V.D. Berg, M.P.Friedlander, “Probing the Pareto Frontier for Basis Pursuit Solutions,” *SIAM Journal on Scientific Computing*, pp.890-912, 2008.