

3次元音響再生のための頭部伝達関数の推定と  
再生制御器の設計に関する研究

2016年度

松井 健太郎

学位論文 博士（工学）

3次元音響再生のための頭部伝達関数の推定と  
再生制御器の設計に関する研究

2016年度

慶應義塾大学大学院理工学研究科

松井 健太郎



# 目次

第 1 章	序論	1
1.1	背景	1
1.2	本論文の目的	4
1.3	本論文の構成	5
第 2 章	研究の背景	9
2.1	頭部伝達関数	9
2.2	頭部伝達関数の測定	11
2.2.1	インパルス応答の測定	11
2.2.2	測定信号	13
2.3	バイノーラル再生法	16
2.4	トランスオーラル再生法	16
2.5	頭部伝達関数に係る諸課題	21
2.5.1	頭部伝達関数の測定	21
2.5.2	バイノーラル再生法	22
2.5.3	トランスオーラル再生法	23
第 3 章	システム同定理論に基づく頭部伝達関数の多方向同時推定	27
3.1	予測誤差法による頭部伝達関数の同時推定	27
3.2	同定入力を作成	29
3.3	実験 1: 頭部伝達関数の多方向同時推定実験	31
3.3.1	実験条件	31
3.3.2	実験結果	34
3.4	実験 2: 提案する同定入力を用いた多方向同時推定実験	40
3.4.1	実験条件	41
3.4.2	実験結果	42
3.5	実験の考察	42

---

第 4 章	出力追従制御を応用したトランスオーラル再生制御器の設計	47
4.1	出力追従制御を応用したトランスオーラル再生法 . . . . .	47
4.2	線形行列不等式 . . . . .	49
4.3	非干渉化制御による逆システムの構成 . . . . .	49
4.4	出力追従制御器の設計手続き . . . . .	53
第 5 章	22.2 マルチチャンネル音響のトランスオーラル再生	63
5.1	22.2 マルチチャンネル音響 . . . . .	63
5.1.1	22.2 マルチチャンネル音響方式の概要 . . . . .	63
5.1.2	22.2 マルチチャンネル音響の家庭での再生 . . . . .	65
5.2	ディスプレイ体型スピーカ . . . . .	66
5.2.1	枠型スピーカアレイ . . . . .	67
5.2.2	ラインアレイスピーカ . . . . .	69
5.2.3	22.2 マルチチャンネル音響トランスオーラル再生 . . . . .	70
5.3	実験 1: 枠型スピーカアレイによるシミュレーション実験 . . . . .	72
5.3.1	実験条件 . . . . .	72
5.3.2	実験結果 . . . . .	76
5.4	実験 2: ラインアレイスピーカによるシミュレーション実験 . . . . .	82
5.4.1	実験条件 . . . . .	83
5.4.2	実験結果 . . . . .	84
5.5	実験の考察 . . . . .	88
第 6 章	結論	91
6.1	研究の成果 . . . . .	91
6.2	今後の発展性 . . . . .	92
参考文献		97
付録 A	特異値分解法によるインパルス応答からの実現	105
付録 B	研究業績	109

# 目次

1.1	Amplitude panning . . . . .	2
2.1	Example of an HRTF . . . . .	10
2.2	Example of an HRIR . . . . .	11
2.3	Binaural microphone . . . . .	12
2.4	HATS (Head and Torso Simulator) . . . . .	13
2.5	Measurement of impulse responses . . . . .	14
2.6	Optimized Aoshima's Time-Stretched Pulse . . . . .	14
2.7	Logarithmic Time-Stretched Pulse . . . . .	15
2.8	Schematic diagram illustrating binaural reproduction . . . . .	17
2.9	Schematic diagram illustrating transaural reproduction . . . . .	18
2.10	Schematic diagram illustrating multichannel transaural reproduction . . . . .	19
3.1	Schematic diagram illustrating simultaneous estimation of HRTFs . . . . .	28
3.2	Block diagram illustrating simultaneous estimation of HRTFs . . . . .	28
3.3	Measurement of impulse responses for simultaneous estimation . . . . .	32
3.4	Loudspeaker arrangement . . . . .	33
3.5	Frequency characteristics of estimated HRIRs (horizontal plane) . . . . .	35
3.5	Frequency characteristics of estimated HRIRs (horizontal plane) . . . . .	36
3.6	Frequency characteristics of estimated HRIRs (median plane) . . . . .	37
3.6	Frequency characteristics of estimated HRIRs (median plane) . . . . .	38
3.7	SDs of estimated HRTFs (horizontal plane) . . . . .	39
3.8	SDs of estimated HRTFs (median plane) . . . . .	39
3.9	Distribution of average SDs of 20 subjects . . . . .	40
3.10	Frequency characteristics of estimated HRIRs . . . . .	43
3.10	Frequency characteristics of estimated HRIRs . . . . .	44
3.11	SDs of estimated HRTFs . . . . .	45

4.1	Schematic diagram illustrating transaural reproduction . . . . .	48
4.2	Design procedure of a tracking controller . . . . .	53
4.3	Parallel structured system for system identification . . . . .	54
4.4	Output tracking controller as a transaural reproduction controller . . . . .	61
5.1	Audio channel arrangement of 22.2 multichannel sound system . . . . .	66
5.2	Concept images of viewing the 8K SHV in home . . . . .	68
5.3	Concept images of viewing the 8K SHV with mobile terminals . . . . .	69
5.4	72-loudspeaker frame jointly developed by NHK and Foster Electric Company, Ltd. . . . .	69
5.5	12-loudspeaker frame integrated into an 85-inch LCD jointly developed by NHK and Foster Electric Company, Ltd. . . . .	70
5.6	Corrugated design of diaphragm . . . . .	70
5.7	Digital loudspeaker frame jointly developed by NHK and Sharp Corp. . . . .	71
5.8	Digital line array loudspeaker jointly developed by NHK and Sharp Corp. . . . .	72
5.9	Layout of sound sources and listening position . . . . .	73
5.10	Measurement of the reproduced sound field in an acoustic anechoic chamber of the NHK Science and Technology Research Laboratories . . . . .	74
5.11	Layout of the loudspeaker units in experiment 1 . . . . .	75
5.12	Comparison of frequency characteristics (simultaneous estimation, red solid line: measured impulse responses, blue solid line: target model) . . . . .	77
5.13	Comparison of frequency characteristics (direction-by-direction measurement, red solid line: measured impulse responses, blue solid line: target model) . . . . .	78
5.14	Comparison of frequency characteristics (simultaneous estimation, red solid line: target model, blue solid line: cascade structured system) . . . . .	79
5.15	Comparison of frequency characteristics (direction-by-direction measurement, red solid line: target model, blue solid line: cascade structured system) . . . . .	80
5.16	Comparison of frequency characteristics (simultaneous estimation, red solid line: measured, blue solid line: convolved) . . . . .	81
5.17	Comparison of frequency characteristics (direction-by-direction measurement, red solid line: measured, blue solid line: convolved) . . . . .	82
5.18	Layout of the loudspeaker units in experiment 2 . . . . .	84
5.19	Comparison of frequency characteristics (red solid line: measured, blue solid line: target model) . . . . .	85

5.20 Comparison of frequency characteristics (red solidline: target model, blue solid line: cascade structured system) . . . . . 86

5.21 Comparison of frequency characteristics (red solid line: measured, blue solid line: synthesized) . . . . . 88



# 表目次

1.1	Nomenclature . . . . .	6
1.2	Important abbreviations . . . . .	7
3.1	Experimental conditions in experiment 1 . . . . .	32
3.2	Input signals in experiment 1 . . . . .	33
3.3	Experimental conditions in experiment 2 . . . . .	41
3.4	Input signals in experiment 2 . . . . .	41
5.1	Channel number, label, and name of 22.2 multichannel sound system . . . . .	67
5.2	Measurement signals in experiment 1 . . . . .	73
5.3	SDs generated by modeling error (simultaneous estimation, dB) . . . . .	76
5.4	SDs generated by modeling error (direction-by-direction measurement, dB) . . . . .	77
5.5	$\mathcal{H}_\infty$ norm of the tracking controllers (simultaneous estimation) . . . . .	79
5.6	$\mathcal{H}_\infty$ norm of the tracking controllers (direction-by-direction measurement) . . . . .	80
5.7	SDs at control points (simultaneous estimation, dB) . . . . .	81
5.8	SDs at control points (direction-by-direction measurement, dB) . . . . .	82
5.9	Measurement signals in experiment 2 . . . . .	83
5.10	SDs generated by modeling error (dB) . . . . .	85
5.11	$\mathcal{H}_\infty$ norm of the tracking controllers . . . . .	86
5.12	Increase of SDs (dB) . . . . .	87
5.13	SDs at control points (dB) . . . . .	88

# 第 1 章

## 序論

### 1.1 背景

コンサートホールやスタジアム，映画やドラマのシーンなど，様々な音場の臨場感を再現，創造する音場再現・再生技術の研究は，古くより音響研究者の興味の一つである．2 台のスピーカによるステレオフォニック（以下，ステレオ）再生から始まったその研究は，前後感，左右感といった平面上の 2 次元的な音響表現から，上下感も含めた 3 次元的な音響表現へと拡張されてきた．近年，その技術は，物理音響モデルに基づく方式と，心理音響モデルに基づく方式とに大別される [1]．

物理音響モデルに基づく方式は，原音場での音響的な物理量を再生音場において忠実に再現しようというものである．その一つである境界音場制御は，ホイヘンスの原理を理論基盤とし，受聴位置を取り囲む閉曲面上の音圧と法線方向の粒子速度を制御し原音場と一致させることにより，閉曲面内に原音場と同一の音場を再現する [1, 2]．この閉曲面を無限大の平面とすると，閉曲面の内外はこの平面を境界とする二つの領域とみなすことができる．この平面上での音圧を原音場と一致させることにより再生音場を制御する方法が波面合成法である [3]．また，音圧の 3 次元的な方向分布を球面調和関数の和として展開，解析し，受聴点に入射する音の方向性を再現するアンビソニックスも，物理音響モデルに基づく方法の一つと考えることができる [4]．

一方，心理音響モデルに基づく方式は，人の音像<sup>\*1</sup>定位知覚に係る心理・神経生理学的な特性を利用したものである [1, 5]．2 チャンネルステレオとその拡張に基づくチャンネルベース音響はその範疇に含まれ，本論文で提案するトランスオーラル再生法は，このチャンネルベース音響を制御の対象としている．

2 チャンネルステレオは，受聴者の前方に配置された L (Left) と R (Right) の 2 台の

---

\*1 距離と方向をもって知覚される感覚上の音源，音の像．

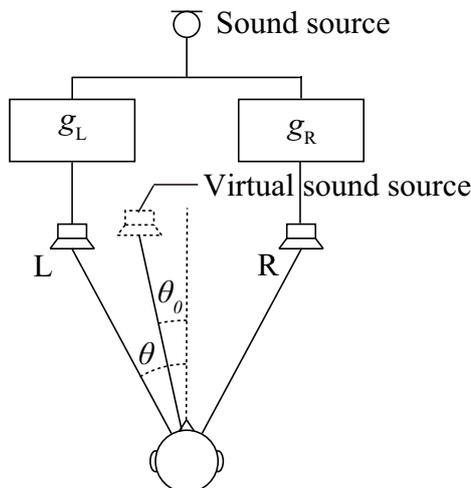


Fig. 1.1: Amplitude panning

スピーカを用いて、そのスピーカ間での音像の広がり感や方向感、定位感を表現するものである。Fig. 1.1 に示すように、同一の音源信号を L チャンネル、R チャンネルから異なる音圧レベルで再生する振幅パンニングを基礎技術としており、その音圧レベル差により受聴者の両耳位置で音源信号のレベル差と位相差を生じさせ、スピーカ間で音の方向感や定位感を操作する。このようにしてスピーカ間に合成される音像を仮想音像という。L、R のスピーカは、受聴者の耳の高さ、正面方向を基準として、左右におよそ  $\theta = 30^\circ$  の間隔をおいて配置される。このとき、仮想音像の方向を  $\theta_0$  とすると、L チャンネル、R チャンネルのゲイン  $g_L, g_R$  は

$$\frac{g_L - g_R}{g_L + g_R} = \frac{\sin \theta_0}{\sin \theta} \quad (1.1)$$

あるいは

$$\frac{g_L - g_R}{g_L + g_R} = \frac{\tan \theta_0}{\tan \theta} \quad (1.2)$$

により与えられる。式 (1.1), (1.2) は振幅パンニングの代表的な計算式であり、前者をサイン則、後者をタンジェント則という。

5.1 サラウンドは、2 チャンネルステレオの L チャンネル、R チャンネルに加え、C (Center) チャンネル、SL (Surround Left) チャンネル、SR (Surround Right) チャンネルの計 5 チャンネル、そして低音効果 (Low Frequency Effects: LFE) 用の LFE チャンネルから構成される。LFE チャンネルはその再生周波数帯域から 0.1 チャンネルとカウント、表記される。C チャンネルは前方音像の安定化、特に映像を伴う場合にはその映像との同期を目的として追加され、SL チャンネル、SR チャンネルは後方からの広がり感、包まれ感を実現するために追加された。2 チャンネルステレオでは表現が困難であった左右方向への音の移動感や広がり感、そしてサラウンドチャンネルによる包まれ感を表現するもの

である。5.1 サラウンドでも 2 チャンネルステレオと同様に、振幅パンニングにより音像の方向感や定位感を操作する。ただし、SL チャンネル、SR チャンネルは前方の L チャンネル、R チャンネル、C チャンネルと比較して疎に配置されているため、音像の安定した定位は困難である。

2 チャンネルステレオや 5.1 サラウンドは、受聴者の両耳を含む水平面上にスピーカを配置した、いわば 2 次元音響方式である。これに対して、水平面外にもスピーカを配置した 3 次元音響方式が近年盛んに検討されている。例えば Auro Technologies 社は、受聴者の頭上、上方を含む 3 レイヤーからなる 9.1, 10.1, 11.1, 13.1 サラウンド Auro-3D を提案、開発している。これらの音響方式では水平面に加え 3 次元的な音の到来や包まれ感を表現することが可能であり、振幅パンニングから拡張された VBAP 法や 3 次元振幅パンニング技術が用いられている。第 5 章で詳しく述べる 22.2 マルチチャンネル音響（以下、22.2ch 音響）方式は、このチャンネルベース音響の最上位に位置する音響方式である。頭上、上方に加え、受聴者の下方にもスピーカを配置し、3 次元的な音の広がり感や方向感、定位感を表現する。既存のマルチチャンネル音響方式との互換性を有し、2018 年に開始が予定されている 8K スーパーハイビジョン（以下、8K SHV）の音響方式となる。

このように、2 チャンネルステレオから始まるチャンネルベース音響は 2 次元音響から 3 次元音響へと発展してきたが、その表現力とシステム規模の拡大とに反比例して、一般家庭への普及は敷居が高くなっている。敷居の一つはスピーカの数であり、推奨されるその配置である。2 チャンネルステレオでは前方に 2 台だったスピーカは、5.1 サラウンドでは後方に、7.1 サラウンドでは側方にサラウンドスピーカが追加された。その配置は国際電気通信連合無線通信部門（International Telecommunication Union - Radiocommunication Sector: ITU-R）の勧告に規定されており、そこから外れた配置では、音場の臨場感、音像の移動感、定位感が低下することが示されている。それに対して、日本の一般家庭の部屋やリビングルームは、マルチチャンネル音響を構成するには必ずしも適切とはいえない広さとレイアウトである。日本オーディオ協会による調査でも、マルチチャンネル音響環境を構築している家庭の多くで勧告基準から外れた配置となっていることが報告されている [6, 7]。また、スピーカケーブルや電源ケーブルの配線、上方や下方へのスピーカの取り付けの煩雑さも、その敷居を高くしている。

ライフスタイルの多様化も課題を複雑にしている。映像や音響コンテンツの視聴スタイルは様々に細分化されてきており、それらを再生するデバイスもパソコンから携帯端末まで多様化が進んでいる。これらのデバイスはその簡便性と利便性ゆえに広く普及しているが、必ずしもマルチチャンネル音響を想定したものではないことが多い。

## 1.2 本論文の目的

本論文の目的は、背景で述べたような視聴スタイルの変化に対し、より少ないスピーカで、より簡易にマルチチャンネル音響を視聴することができる再生法を提示することである。本論文で提案するトランスオーラル再生法は、人の聴覚特性を利用し、音像を任意の方向に定位知覚させる技術である。この技術を援用して、限定されたスピーカ配置によりマルチチャンネル音響を擬似再生する。また、トランスオーラル再生法を実現するためには、上記の聴覚特性を表現する頭部伝達関数 (Head-Related Transfer Function: HRTF) を多方向にわたって測定する必要がある。その測定負荷を軽減するための多方向同時推定法も提案する。

音響工学では、音響エコーキャンセラや振動制御、アクティブノイズコントロール (Active Noise Control: ANC) などのごく限られた研究分野を除いて、現代制御以降の制御理論をあまり積極的に取り入れてこなかった。これは、人の可聴域を制御対象としたときその周波数帯域幅が非常に広がること、この周波数帯域幅をカバーするサンプリング周波数が高くなることが主な理由として考えられる。高い周波数でのリアルタイム処理や適応制御は、その制御器に高速な演算能力と広い作業領域を要求する。特に音場再現・再生では多数の 2 次音源を用いての多チャンネル制御が主流となっており、このチャンネル数の増加により演算量も加速度的に増加する。上記の ANC においても、著者の知る限り演算の高速化や演算量の低減を志向する研究が多く、組み合わせるパッシブ制御との兼ね合いもあるが、その制御対象となる騒音も中・低周波数帯域に限定されるものがほとんどであった。

しかし近年、汎用の DSP (Digital Signal Processor) や FPGA (Field-Programmable Gate Array) が普及し、パーソナルコンピュータに搭載される CPU (Central Processing Unit) のマルチコア化、作業領域となるメモリの低廉化が進んだ結果、特別な環境を整備しなくとも十分な規模での高速演算が可能となった。また、多数のチャンネルを扱うための入出力フォーマットやインターフェースが整備され、それらのフォーマットをサポートする電気音響機器も増えている。こうした環境変化のもと、音場再現・再生や音源推定、3 次元收音などの様々な研究分野において、適応制御やロバスト制御を応用した新しい研究アプローチが見られるようになってきた [8–11]。

本論文で提案するトランスオーラル再生法はもまた、その制御器をロバスト制御理論に基づいて設計する。音場再現・再生技術の新しい研究アプローチとしてその進捗を示すことも、本論文の目的の一つである。

## 1.3 本論文の構成

本論文の構成は以下のとおりである。

第 2 章では、本論文の基礎となる HRTF について、その定義から測定、応用となるバイノーラル再生法とトランスオーラル再生法までを概説する。

第 3 章では、HRTF の測定時間を短縮する方法として、システム同定理論に基づく多方向同時推定法を提案する。また、同定入力を一つの  $M$  系列 (Maximum Length Sequence: MLS) から作成する方法も示す。提案法では、HRTF の集合を多入力 1 出力 (Multi Input Single Output: MISO) の有限インパルス応答 (Finite Impulse Response: FIR) モデルとみなし、予測誤差法によりそのパラメータを推定する。1 方向ずつの測定と比較することにより、提案法の有効性を示す。

第 4 章では、出力追従制御を応用した状態空間でのトランスオーラル再生法を提案する。トランスオーラル再生法は、人の聴覚特性を利用し、音像を任意の方向に定位知覚させる技術である。提案法では、トランスオーラル再生制御器の設計を、制御器の  $\mathcal{H}_\infty$  ノルムを評価指標とする凸計画問題として定式化する。制約の記述には線形行列不等式 (Linear Matrix Inequality: LMI) を用いる。トランスオーラル再生制御器の設計に係る制御仕様が、適当な等価変換と変数変換により LMI に帰着されることを示す。

第 5 章では、チャンネルベースの 3 次元音響方式である 22.2ch 音響方式を概説したのち、提案するトランスオーラル再生法の適用について述べる。22.2ch 音響は多様な視聴スタイルが予想されるが、多くの場合、家庭で 24 台のスピーカを配置することは困難と考えられる。本章ではまず、ディスプレイに一体化される枠型スピーカアレイとラインアレイスピーカを用いた 22.2ch 音響のトランスオーラル再生法について述べる。次に、同スピーカを用いたシミュレーション実験により、提案する HRTF の多方向同時推定法とトランスオーラル再生法の有効性を示す。

第 6 章では、本論文をまとめる。

本論文で用いる記号を Table 1.1 にまとめる。ただし、サイズは必要に応じて記すことにする。また、本論文で用いる略記を Table 1.2 にまとめる。なお、本論文では離散時間線形時不変システムを対象とする。

Table 1.1: Nomenclature

---

Notation	Description
$\mathbb{R}^n$	The set of real vectors of size $n$
$\mathbb{R}^{m \times n}$	The set of real matrices of size $m \times n$
$\mathbb{C}^n$	The set of complex vectors of size $n$
$\mathbb{C}^{m \times n}$	The set of complex matrices of size $m \times n$
$\mathbb{S}_{++}^n$	The set of positive definite matrices of size $n \times n$
$\mathbf{0}_{m \times n}$	Zero matrix of size $m \times n$
$\mathbf{I}_n$	Identity matrix of size $n$
$z$	Any complex number working as phase shift in the $z$ -domain

---

Table 1.2: Important abbreviations

Abbreviation	Full name
ANC	Active Noise Control
ARIB	Association of Radio Industries and Businesses
FPD	Flat Panel Display
FIR	Finite Impulse Response
HATS	Head and Torso Simulator
HRIR	Head-Related Impulse Response
HRTF	Head-Related Transfer Function
IID	Interaural Intensity Difference
ILD	Interaural Level Difference
IPD	Interaural Phase Difference
ITD	Interaural Time Difference
ITU-R	International Telecommunication Union - Radiocommunication Sector
LFE	Low Frequency Effects
LMI	Linear Matrix Inequality
Log-TSP	Logarithmic Time-Stretched Pulse
MIMO	Multi Input Multi Output
MISO	Multi Input Single Output
MLS	Maximum Length Sequence
OATSP	Optimized Aoshima's Time-Stretched Pulse
SD	Spectral Distortion
SDP	Semidefinite Programming Problem
SMPTE	Society of Motion Picture and Television Engineers
TSP	Time-Stretched Pulse
TSVD	Truncated Singular Value Decomposition
YALMIP	Yet Another LMI Parser



## 第 2 章

# 研究の背景

本論文で提案するトランスオーラル再生法は，人の聴覚特性を利用し，音像を任意の方向に定位知覚させる技術である．本章では，このトランスオーラル再生法とその基礎となる頭部伝達関数 (Head-Related Transfer Function: HRTF) について，その定義から測定法，応用までを概説する．

### 2.1 頭部伝達関数

音像の定位知覚に係る最も信頼できる手掛かりは，両耳に到達する音の比較である [12]．頭部の一方側にある音源から発する音は，左右の耳に入射する際に，その経路差により到達時間と音圧に差を生じる．すなわち，遠い方の耳に入射する音は近い方の耳に入射する音に比較して，時間的に遅れ，音圧は低くなる．この時間差を両耳間時間差 (Interaural Time Difference: ITD) といい，音圧差を両耳間レベル差 (Interaural Level Difference: ILD) という<sup>\*1</sup>．人は，この ITD と ILD を手掛かりとして，主に水平面上での音像の定位を知覚している．それぞれの寄与は音の性質によって異なってくるが，例えば定常的な正弦波音の場合，頭部の大きさに対して波長の長い低周波数帯域では回折の結果として ILD が小さくなり，ITD の寄与が大きくなる．逆におおよそ 1,500 Hz を境として，高周波数帯域では ILD が重要となってくる [12]．

一方，仰角方向，とりわけ両耳間の音に差のほとんどない正中面において，音像の定位知覚に重要な役割を果たすのが頭部や耳介である [12–14]．左右の耳に入射する音は，頭部に対する入射方向によりそのスペクトルを変化させる．この変化は頭部や胴部における反射や回折，耳介における共振に起因するものであり，入射方向に依存して特徴的なスペクトルパターンを形成する．人は，このスペクトルパターンと音の入射方向との関係を学

---

<sup>\*1</sup> ITD, ILD はそれぞれ両耳間位相差 (Interaural Phase Difference: IPD), 両耳間強度差 (Interaural Intensity Difference: IID) ともいわれる．

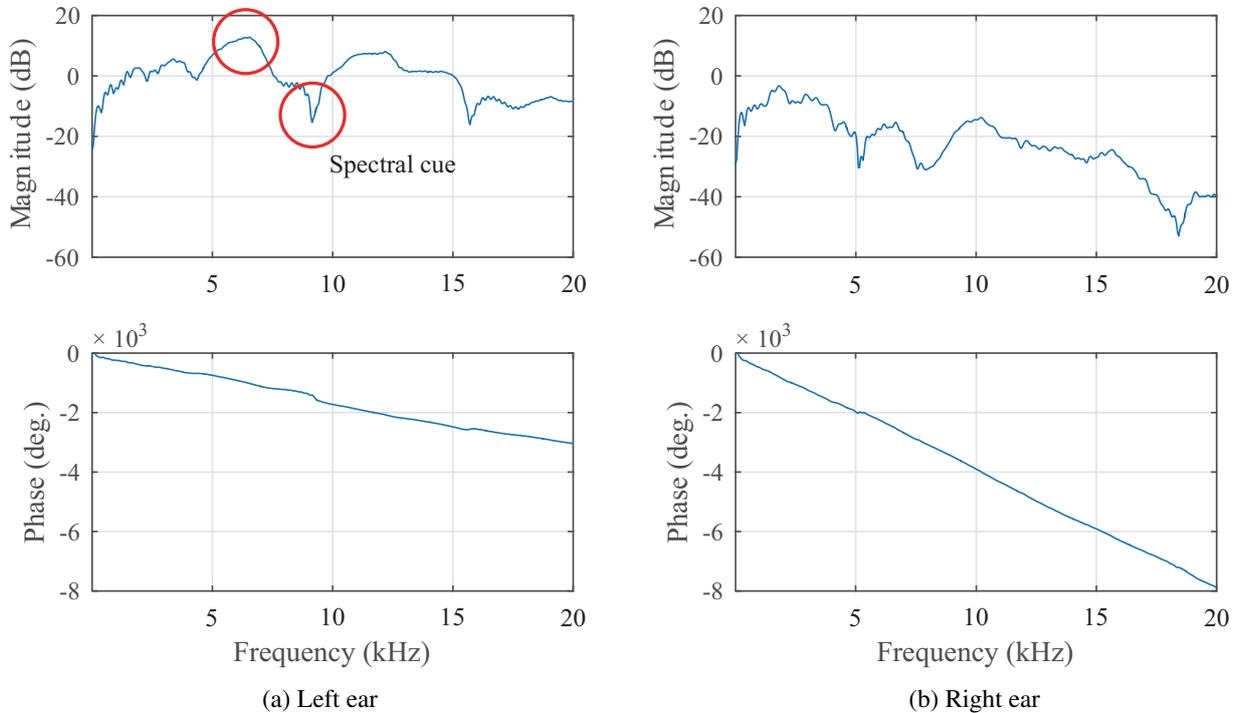


Fig. 2.1: Example of an HRTF

習により獲得し、仰角方向の定位知覚の手掛かりの一つとしている。

これら ITD や ILD，入射方向に依存するスペクトルパターンなど，音像の定位知覚に係る手掛かりを包含する伝達関数が HRTF である。HRTF は，自由音場において，頭のない状態での頭部中心に相当する両耳中心位置から，頭のある状態での両耳鼓膜位置あるいは外耳道入口までの音響伝達関数として，または近似的に，頭外音源位置から両耳鼓膜位置あるいは外耳道入口までの音響伝達関数として定義される，音源位置と周波数に係る伝達関数である [15]。前述のスペクトルパターンは振幅特性上の特徴的なピークやノッチにより構成され，これらピーク，ノッチはスペクトルキューと呼ばれる。例として，左側方  $90^\circ$  の HRTF を Fig. 2.1 に，その時間領域表現である頭部インパルス応答 (Head-Related Impulse Response: HRIR) を Fig. 2.2 に示す。Fig. 2.1 (a) は左耳の HRTF，(b) は右耳の HRTF である。それぞれ上図は振幅特性，下図は位相特性を表す。振幅特性上に緩やかな山を形成しているのがスペクトルキューのピークであり，比較的急峻な谷を形成しているのがノッチである。スペクトルキューの多くは  $5 \sim 10$  kHz の間に存在する [16]。例えば，4 kHz 付近のピーク (P1) とそのピークよりやや高い周波数帯域に存在するノッチ (N1, N2) は，P1 を基準として N1, N2 の中心周波数が音の入射角度に応じて遷移することか

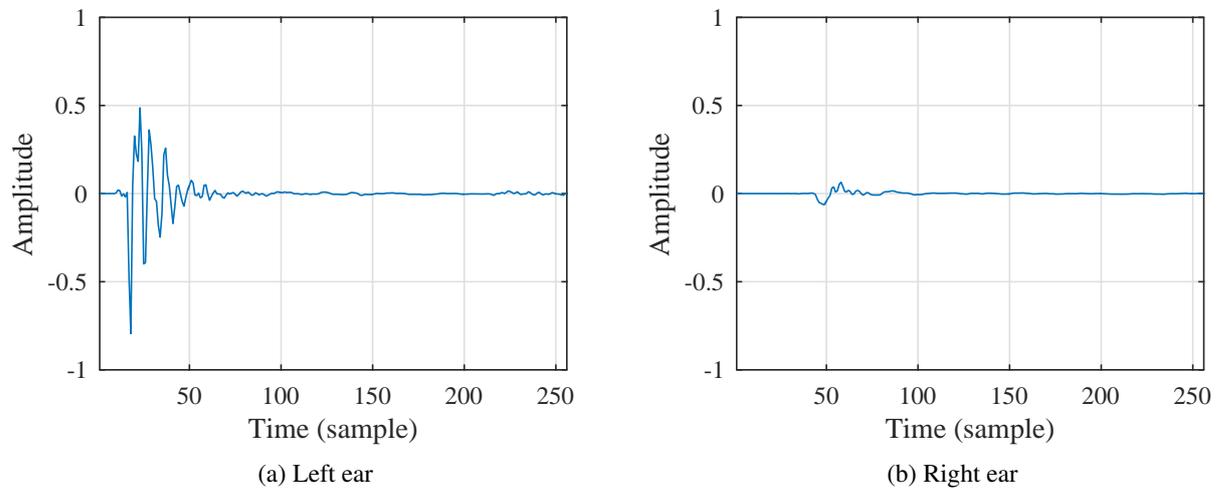


Fig. 2.2: Example of an HRIR

ら，正中面での音像の仰角知覚に重要な役割を果たすとされている [17]．Fig. 2.2 (a) は左耳の HRIR，(b) は右耳の HRIR である．応答の立ち上がり前，無音区間の時間差が ITD であり，応答の振幅差が ILD である．なお，Fig. 2.1 に示す HRTF の測定はサンプリング周波数 48 kHz で実施し，スピーカとマイクロフォンの特性および人の可聴域を考慮して帯域を 20 Hz から 20 kHz に制限した．

## 2.2 頭部伝達関数の測定

HRTF は，その定義より，頭外音源位置から外耳道入口までの音響伝達関数を，音源位置から両耳中心位置までの音響伝達関数で除した伝達関数として与えられる．通常，音源位置から外耳道入口まで，音源位置から両耳中心位置までのインパルス応答をそれぞれ測定し，フーリエ変換により伝達関数に変換したのち，周波数領域で除算することにより HRTF を取得する．

### 2.2.1 インパルス応答の測定

インパルス応答の測定は，実際に人の特性を実測する方法と，HATS (Head and Torso Simulator) やダミーヘッドを使用する方法の 2 種類があり，通常，音響無響室で実施される．

人の特性を実測する場合，その両耳の鼓膜近傍あるいは外耳道入口にマイクロフォンを据え付ける．鼓膜近傍での測定には，主にプローブマイクロフォンが使用される．プロー

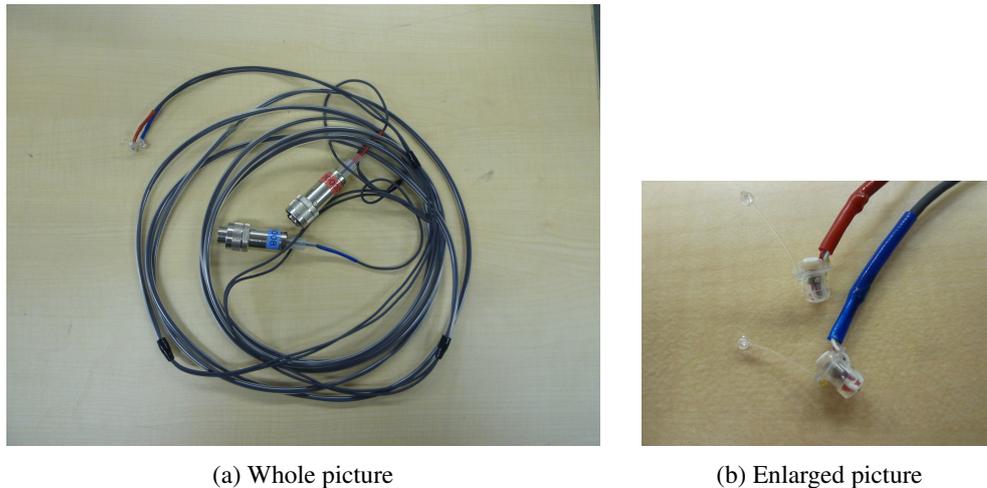


Fig. 2.3: Binaural microphone

ブマイクログフォンはプローブ拡張チューブとマイクログフォンを組み合わせた構成をしており、外耳道のような狭い環境で、近傍の音場をあまり乱すことなく音圧を測定することができる。ただし、プローブ先端を外耳道奥に挿入しての測定となるため、外耳道内や鼓膜を傷つけることがないように安全に注意を払う必要がある。プローブマイクログフォンの据え付けには耳科医の協力を得るべきとの指摘もある [14]。外耳道入口での測定は、鼓膜近傍での測定の代替として実施される。これは、外耳道内の伝達関数が音源位置にはあまり依存しないためである。外耳道入口での測定を目的として開発したバイノーラルマイクログフォンを Fig. 2.3 に示す。カナル型イヤフォンとよく似た形状をしており、イヤピースを挿入したとき、外耳道入口より外を向くように小型のエレクトレットコンデンサマイクログフォンが備え付けられている。(a) は全体図、(b) はマイクログフォンカプセルの拡大図となる。イヤピースは外耳道の大きさに合わせて交換することが可能であり、インパルス応答は外耳道を閉塞した状態で測定される。

HATS はマウスシミュレータとイヤースミュレータを組み込んだマネキンであり、その形状は平均的な男性、女性の頭部と胴部の寸法に基づいて製作されている (Fig. 2.4)。人の特性を実測する場合と同様に頭部や胴部の影響を再現した測定が可能であることから、音響収録や音質評価、被測定者の負担が見込まれる測定などで広く使用されている。通常、頭部と胴部を備えたマネキンを HATS、頭部のみのもをダミーヘッドというが、慣習的に両者を区別することなくダミーヘッドということもある。

Fig. 2.5 は上半球面のインパルス応答測定の様子である。スピーカは被測定者の頭部位置を中心とする円弧上に  $10^\circ$  間隔で配置されている。この距離は通常、スピーカより発せられる音波を平面波とみなして問題のない  $1.2 \sim 1.5$  m 程度とされており、図の測定では  $1.3$  m とした。音源となるスピーカから測定信号を順番に再生し、被測定者の外耳道入口



Fig. 2.4: HATS (Head and Torso Simulator)

に挿入されたバイノーラルマイクロフォンで応答を收音する．被測定者の座る椅子はステッピングモータによる回転機構を備えており，この制御により上半球面のインパルス応答を1方位ずつ測定する．

### 2.2.2 測定信号

測定信号は高い信号対雑音比 (Signal to Noise Ratio: SNR, SN 比) と広帯域の周波数特性を持つ必要があることから，M 系列などの擬似ランダム信号，時間引き延ばしパルス (Time-Stretched Pulse: TSP) [18] などのスイープ信号が使用される．TSP は，単位インパルスの位相を周波数の関数としてシフトすることにより，時間方向に引き延ばした信号である<sup>\*2</sup>．その一例として，OATSP (Optimized Aoshima's Time-Stretched Pulse) [20] の時間波形を Fig. 2.6 に，Log-TSP (Logarithmic Time-Stretched Pulse) [21] の時間波形を Fig. 2.7 に示す．いずれの図も (a) は TSP，(b) はその逆特性を持つ FIR フィルタであり，OATSP，Log-TSP とともに TSP の最大振幅が  $-2$  dB (Full Scale) となるように正規化している．OATSP は，単位インパルスの位相を周波数の2乗に比例してシフトした TSP をもとに，ナイキスト周波数で位相が不連続となる問題を解消するよう修正した信号である．

<sup>\*2</sup> 引き延ばされた TSP はパルスの範囲ではないということで Swept-sine と呼称されることもあるが，本論文では慣習に倣い TSP と呼ぶことにする [19] ．



Fig. 2.5: Measurement of impulse responses in an acoustic anechoic chamber of the NHK Science and Technology Research Laboratories

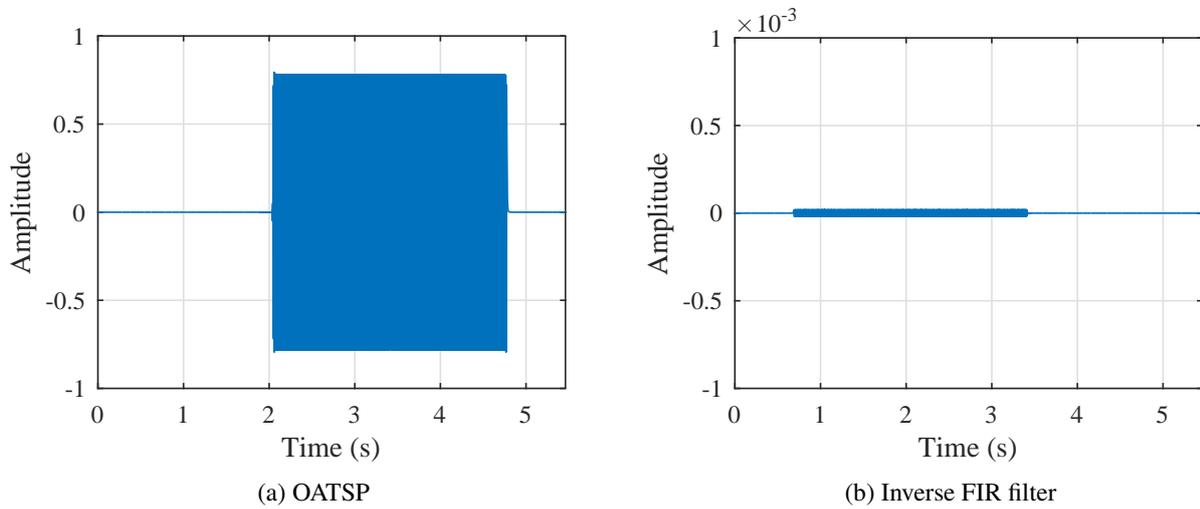


Fig. 2.6: Optimized Aoshima's Time-Stretched Pulse

信号長とスweep区間を決定する整数  $m$  に対して, その周波数特性  $U(n)$  は

$$U(n) = \begin{cases} \exp \frac{j4m\pi n^2}{N^2}, & 0 \leq n \leq \frac{N}{2} \\ U^*(N - n), & \frac{N}{2} < n < N \end{cases} \quad (2.1)$$

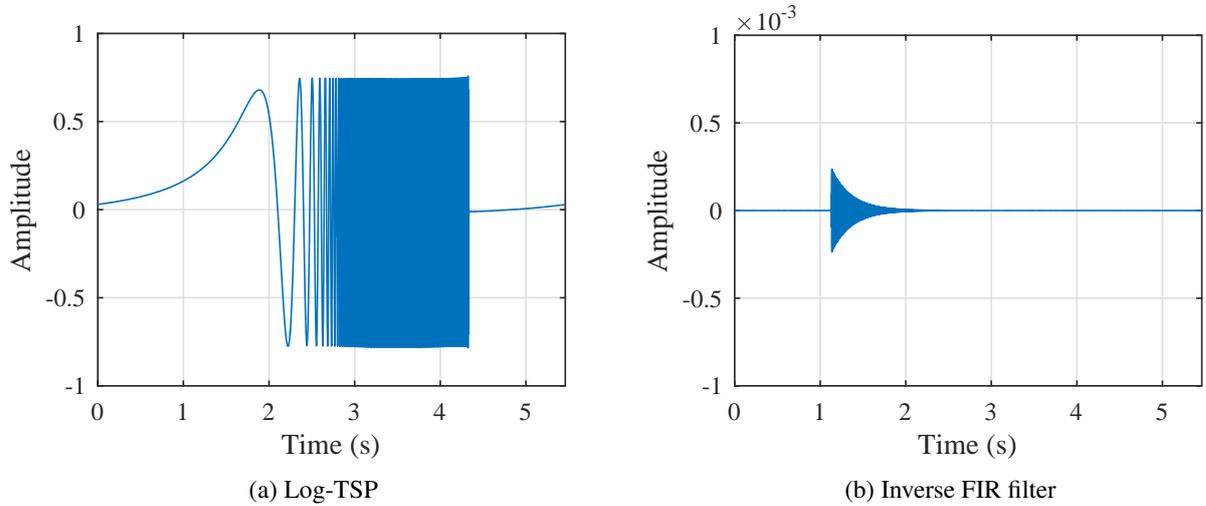


Fig. 2.7: Logarithmic Time-Stretched Pulse

により与えられる．また，Log-TSP は，対数軸上で位相をシフトさせた信号であり，その周波数特性  $U(n)$  は

$$U(n) = \begin{cases} 1, & n = 0 \\ \frac{\exp(ja_0 n \log n)}{\sqrt{n}}, & 0 < n \leq \frac{N}{2} \\ U^*(N - n), & \frac{N}{2} < n < N \end{cases} \quad (2.2)$$

により与えられる．ただし，

$$a_0 = \frac{2m\pi}{(N/2) \log(N/2)} \quad (2.3)$$

である．また，式 (2.1)，(2.2) のいずれにおいても， $n$  は離散周波数ピンを， $N$  はその数を，上付き文字 ‘\*’ は複素共役を表す．OATSP は白色のスペクトルを持つことから，全周波数帯域に渡って等しい SN 比での測定となる．一方，Log-TSP はピンク色のスペクトルを持つため，高周波数帯域に対する低周波数帯域の相対的な SN 比が改善される [21]．また，対数軸上で位相をシフトさせるため，高調波歪みに代表される非線形歪みの分離が可能であるという特長を持つ [22]．いずれの測定信号を使用する場合も，收音された応答に逆特性を持つ FIR フィルタを畳み込むことによりインパルス応答を得ることができる．通常，測定信号を複数回連続して再生し，收音された応答を同期加算することにより SN 比の改善を図る．このとき，直前の測定信号に対する応答の折り返しが重畳されるため，最初と最後の応答を切り捨てた上で，直線畳み込みではなく巡回畳み込みを適用する必要がある．

## 2.3 バイノーラル再生法

2.1 節で述べたように，HRTF には ITD や ILD，振幅特性上のスペクトルキューなど，音像の定位知覚に係る多くの特徴量が含まれている．そのため，HRTF を音源信号に作用させる，すなわち所望方向の左右耳 HRIR を音源信号に畳み込み両耳に直接提示することにより，任意の方向にその音像を定位知覚させることができる．このようにして生成される信号をバイノーラル信号，バイノーラル信号をヘッドフォンやイヤフォンにより提示する再生法をバイノーラル再生法という．また，バイノーラルマイクロフォンや HATS を使って外耳道位置での信号を收音することをバイノーラル收音といい，その信号を両耳に直接提示する再生法もまたバイノーラル再生法という．

Fig. 2.8 に典型的なバイノーラル再生の概略図を示す．バイノーラル再生法および 2.4 節で説明するトランスオーラル再生は，時間領域，周波数領域， $z$  領域など様々な領域での処理に基づく再生法が提案されているが，ここでは概略説明のため領域を表す引数は省略した．図中， $G^l, G^r$  は左右耳の HRTF を表す．(a) は HATS を使って收音したバイノーラル信号をヘッドフォンにより両耳に直接提示する方法であり，(b) は音源信号に左右耳の HRTF を作用させた信号をバイノーラル信号として，同側の耳に直接提示する方法である．ヘッドフォンで受聴する場合，ハウジング内での反射や共振の影響により信号の周波数特性が変化するため，その特性を逆フィルタにより補償する必要がある． $X^l, X^r$  はそのための補償フィルタである．

## 2.4 トランスオーラル再生法

バイノーラル信号をスピーカより提示することを考える．両耳に直接信号を提示することができるヘッドフォン受聴と異なり，スピーカ受聴の場合には，スピーカから同側耳への信号の伝搬に加え，対側耳への信号の漏えい(クロストーク)が発生する．したがって，このクロストークを抑制し，所望信号のみをそれぞれの耳に伝送するための補償処理が必要となる．この補償処理をクロストーク・キャンセレーションといい，クロストーク・キャンセレーションにより実現されるスピーカによるバイノーラル再生法を，ヘッドフォンによる再生法と区別してトランスオーラル再生法という．

クロストーク・キャンセレーションに基づくスピーカ再生法は，Bauer [23] によって提案され，後に Atal, Schroeder [24–26] の手によって定式化された．Schroeder はマネキンを使って測定したインパルス応答からクロストーク・キャンセラを設計し，各所のコンサートホールで実施したバイノーラル録音の音響無響室でのスピーカ再生を試みている．Schroeder らの解析的な方法に対し，Damaske ら [27, 28] は，左右のスピーカに印加され

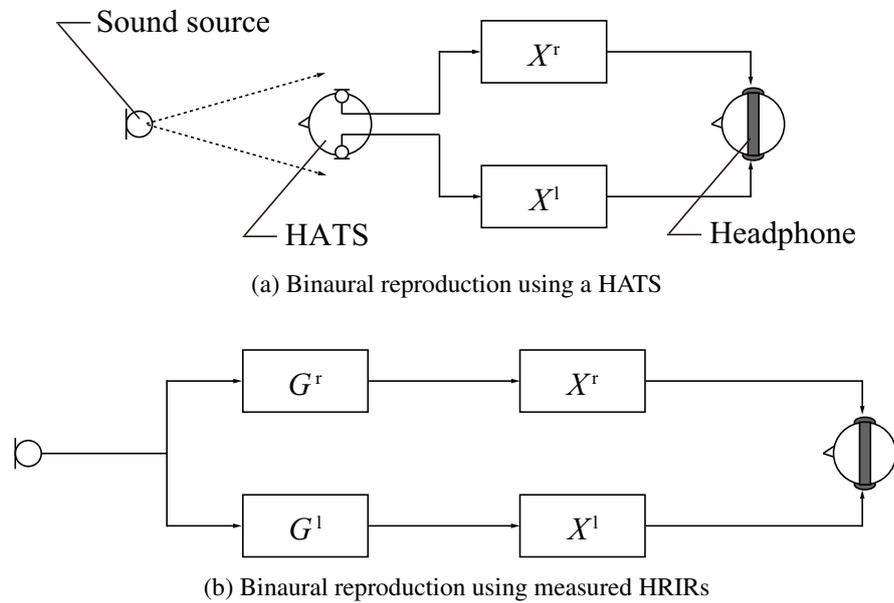


Fig. 2.8: Schematic diagram illustrating binaural reproduction

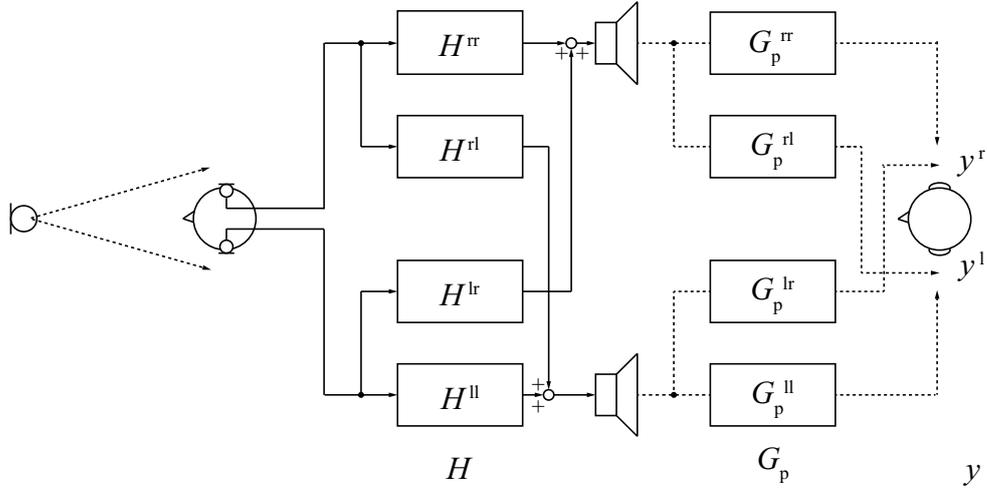
る信号の遅延量と減衰量を試行錯誤により調整し，実験的にクロストーク・キャンセラを設計する方法を提案した．Cooper, Bauck [29, 30] は，身体の左右対称性を仮定してモデル化することにより，クロストーク・キャンセラの構造と処理手続きを簡単化した．逆に安藤ら [31] は，頭部の非対称性を考慮したクロストーク・キャンセラの設計を提案した．これら初期の研究は，主に受聴者の前方に配置された 2 台のスピーカによる制御を対象としているが，現在のトランスオーラル再生法の基本形となっている．

Fig. 2.9 に，2 台のスピーカによる典型的なトランスオーラル再生の概略図を示す．ただし，概略説明のため領域を表す引数は省略した．図中， $G_p$  はそれぞれのスピーカから左右耳までの HRTEF を， $G_t$  は左右耳に提示する所望方向の HRTEF を， $H$  はクロストーク・キャンセレーションのための制御器をそれぞれ表す．また，制御対象となる受聴者の耳の位置を制御点，制御音源となるスピーカを 2 次音源という．このとき，入力信号  $u$  から出力信号  $y$  までは

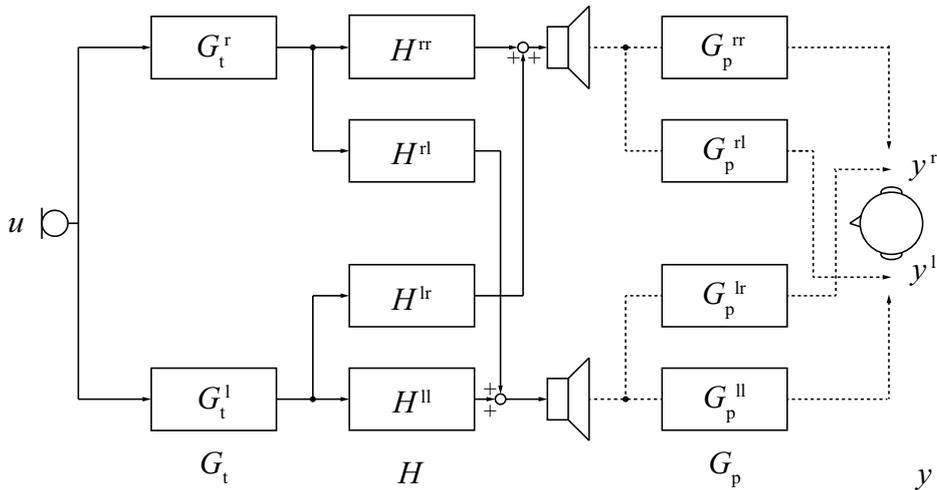
$$y = G_p H G_t u \quad (2.4)$$

と表される．ただし，

$$y = \begin{bmatrix} y^l \\ y^r \end{bmatrix} \quad (2.5)$$



(a) Transaural reproduction using a HATS



(b) Transaural reproduction using measured HRIRs

Fig. 2.9: Schematic diagram illustrating transaural reproduction

であり,

$$\mathbf{G}_p = \begin{bmatrix} G_p^{ll} & G_p^{rl} \\ G_p^{lr} & G_p^{rr} \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} H^{ll} & H^{rl} \\ H^{lr} & H^{rr} \end{bmatrix}, \quad \mathbf{G}_t = \begin{bmatrix} G_t^l \\ G_t^r \end{bmatrix} \quad (2.6)$$

である.

トランスオーラル再生法では, 制御点において所望信号, すなわち入力信号に所望方向のHRTFを作用させた信号が合成されるように設計される. これを式で表すと

$$\mathbf{y} = \mathbf{G}_t \mathbf{u} \quad (2.7)$$

となる. したがって, 式(2.4), (2.7)の関係より, 制御器  $\mathbf{H}$  は  $\mathbf{G}_p$  の逆システムとして設

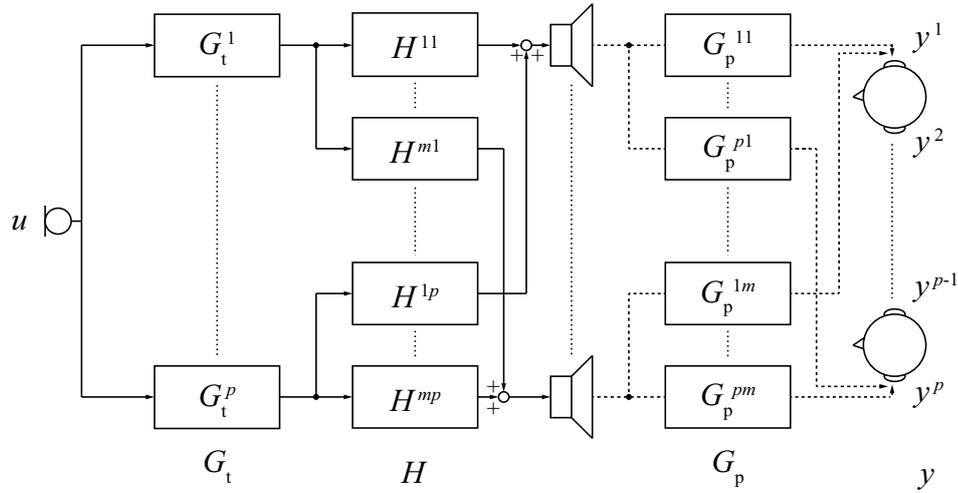


Fig. 2.10: Schematic diagram illustrating multichannel transaural reproduction

計される .

$$\mathbf{H} = \mathbf{G}_p^{-1} \quad (2.8)$$

この制御器の設計に関しては、様々なアプローチが検討されている。例えば周波数伝達関数に基づく設計では、式 (2.8) は離散周波数ビンごとの逆行列計算となる。

$$\mathbf{H}(n) = \mathbf{G}_p^{-1}(n), \quad 1 \leq n \leq N \quad (2.9)$$

ただし、

$$\mathbf{G}_p(n) = \begin{bmatrix} G_p^{ll}(n) & G_p^{rl}(n) \\ G_p^{lr}(n) & G_p^{rr}(n) \end{bmatrix} \in \mathbb{C}^{2 \times 2} \quad (2.10)$$

$$\mathbf{H}(n) = \begin{bmatrix} H^{ll}(n) & H^{rl}(n) \\ H^{lr}(n) & H^{rr}(n) \end{bmatrix} \in \mathbb{C}^{2 \times 2} \quad (2.11)$$

であり、 $n$  は離散周波数ビンを、 $N$  はその数を表す。すべての離散周波数ビンに対して式 (2.9) を計算したのち、時間領域に変換し FIR フィルタとして実装する。

$$\mathbf{h}(k) = \mathcal{F}^{-1} [\mathbf{H}(n)] \quad (2.12)$$

ただし、 $\mathcal{F}^{-1} [\cdot]$  は逆フーリエ変換を表す。

Fig. 2.9 は、左右耳の 2 点を制御点とする、2 台のスピーカによるトランスオーラル再生を示しているが、これは 3 台以上の複数スピーカを用いた方法に容易に拡張することができる。Fig. 2.10 に、複数台のスピーカによるトランスオーラル再生の概略図を示す。図

中,  $p$  は制御点の数を,  $m$  は2次音源の数を表す. このとき, 式 (2.10), (2.11) は

$$\mathbf{G}_p(n) = \begin{bmatrix} G_p^{11}(n) & G_p^{12}(n) & \dots & G_p^{1m}(n) \\ G_p^{21}(n) & G_p^{22}(n) & \dots & G_p^{2m}(n) \\ \vdots & \vdots & \ddots & \vdots \\ G_p^{p1}(n) & G_p^{p2}(n) & \dots & G_p^{pm}(n) \end{bmatrix} \in \mathbb{C}^{p \times m} \quad (2.13)$$

$$\mathbf{H}(n) = \begin{bmatrix} H^{11}(n) & H^{12}(n) & \dots & H^{1p}(n) \\ H^{21}(n) & H^{22}(n) & \dots & H^{2p}(n) \\ \vdots & \vdots & \ddots & \vdots \\ H^{m1}(n) & H^{m2}(n) & \dots & H^{mp}(n) \end{bmatrix} \in \mathbb{C}^{m \times p} \quad (2.14)$$

となる. 2次音源数を増やすとその数に比例して制御点の数も増やすことができるため, 複数スピーカによる制御は, 受聴範囲を広げる有力な手段である.  $p = m$  かつランク落ちが生じていないとき  $\mathbf{G}_p(n)$  は正則となるため,  $\mathbf{H}(n)$  は一意に定まる.  $p < m$  のとき式 (2.9) は劣決定問題となり,  $\mathbf{H}(n)$  は不定となる. そのため, 通常, Moore-Penrose の擬似逆行列を式 (2.9) の解として用いる.

擬似逆行列は

$$\mathbf{H}(n) = \mathbf{G}_p^H(n) \left\{ \mathbf{G}_p(n) \mathbf{G}_p^H(n) \right\}^{-1} \quad (2.15)$$

により求められる. ただし, 上付き文字 'H' は複素共役転置を表し, 式 (2.15) が解を持つためには,  $\mathbf{G}_p(n) \mathbf{G}_p^H(n)$  が正則である必要がある. また, 擬似逆行列は特異値分解を用いて以下のように求めることもできる. まず,  $\mathbf{G}_p(n)$  の特異値分解を行う.

$$\mathbf{G}_p(n) = \mathbf{U}(n) \begin{bmatrix} \boldsymbol{\Sigma}(n) & \mathbf{0} \end{bmatrix} \mathbf{V}^H(n), \quad 1 \leq n \leq N \quad (2.16)$$

ただし,  $\mathbf{U}(n) \in \mathbb{C}^{p \times p}$ ,  $\mathbf{V}(n) \in \mathbb{C}^{m \times m}$  はユニタリ行列であり,  $\boldsymbol{\Sigma}(n) \in \mathbb{R}^{p \times p}$  は対角成分に  $\mathbf{G}_p(n)$  の特異値  $\sigma_i(n)$  ( $i = 1, \dots, p$ ) を持つ対角行列である.

$$\boldsymbol{\Sigma}(n) = \text{diag} [\sigma_1(n), \dots, \sigma_p(n)] \quad (2.17)$$

次に,  $\boldsymbol{\Sigma}(n)$  の逆数を対角成分に持つ対角行列  $\boldsymbol{\Gamma}(n) \in \mathbb{R}^{p \times p}$  を計算する.

$$\boldsymbol{\Gamma}(n) = \text{diag} [\gamma_1(n), \dots, \gamma_p(n)] \quad (2.18)$$

ただし,

$$\gamma_i(n) = \begin{cases} \frac{1}{\sigma_i(n)}, & \sigma_i(n) \neq 0 \\ 0, & \text{otherwise} \end{cases} \quad (2.19)$$

である. 式 (2.18) を用いて, 擬似逆行列は

$$\mathbf{H}(n) = \mathbf{V}(n) \begin{bmatrix} \boldsymbol{\Gamma}(n) \\ \mathbf{0} \end{bmatrix} \mathbf{U}^H(n) \quad (2.20)$$

と表すことができる。

以下、2台のスピーカによるトランスオーラル再生と同様に、すべての離散周波数ビンに対して式(2.20)を計算したのち、逆フーリエ変換により時間領域に変換しFIRフィルタとして実装する。

## 2.5 頭部伝達関数に係る諸課題

### 2.5.1 頭部伝達関数の測定

HRTFの測定では、測定環境で混入する雑音の抑制が課題となる。そのために、測定を実施する実験室の暗騒音を下げると共に、実験室内の反射物や遮蔽物をできるだけ排除する工夫が必要である。特に多数のスピーカを使って多方向のHRTFを測定する場合、配置されたスピーカやスタンドなどの測定治具も反射物となる。また、人の特性を実測する場合、測定時に被測定者の姿勢を維持するため背もたれや頭部を支持する保持具が用いられることが多いが、これらの治具も反射物として、その影響を考慮しなければならない。

スピーカやマイクロフォンなどの電気音響変換器や測定に使用する電気音響機器の発する雑音にも注意を払う必要がある。他の機器との干渉により生じるハム、スピーカへの過大入力による音割れや歪み、アンプやスピーカの非線形性に起因する混変調歪みや高調波歪み、ジッタによる位相雑音、そのほかコネクタの接触不良による雑音などがこれに該当する。これらの雑音を抑制するためには、機器のダイナミックレンジや感度を考慮して測定信号のレベルを適切に調整すると共に、ケーブルの防磁加工や機器間の同期の管理など、雑音の種類に応じた対策を施さなければならない[32]。

HRTFの測定においてもう一つ課題となるのが、測定に費やされる時間である。HRTFの測定は通常1方向ずつ行われる。そのため、測定方向数が増加すると、その数に比例して長時間を要するようになり、測定作業が煩雑となる。例えば、人の聴覚における空間分解能に基づいて水平面上のHRTFを測定する場合、その測定方向数は数十から数百となる。Fig. 2.5に示したように、通常、椅子あるいはトラバーサを回転させながらの順次測定となるため、そのインターバルを含めると、測定時間は1時間を超えることもある。

この測定時間を短縮する方法として、Zotkinら[33]は、ヘルムホルツの相反定理に基づく相反法を提案した。これは、被測定者の外耳道入口に小型のスピーカを据え付け、そのスピーカより再生される測定信号を多方向に配置したマイクロフォンで收音する、通常とは再生点と收音点を入れ替えた方法である。理論的には、複数方向のHRTFをその数に制限なく同時に測定することが可能となるが、両耳に据え付けられる小型スピーカの周波数特性や出力音圧、指向特性が検討課題となっている[34–36]。また、Majdakら[37]、Pulkkiら[38]、Fukudomeら[39]はいずれも、スイープ信号を測定信号とするHRTFの

連続測定法を提案した。この方法では、スイープする測定信号を、多方向から一定時間間隔で再生する、あるいは被測定者の座るサーボ回転椅子を一定速度で回転しながら再生する。このとき、スイープする信号の一部をオーバーラップさせることにより測定時間の短縮を図っている。測定時間の短縮とはやや異なるが、西野ら [40–42]、松本ら [43,44] は、ある限られた方向の HRTF を測定し、時間領域、周波数領域での線形補間やスプライン補間、多項式近似により未測定方向の HRTF を内挿する方法を提案した。また、渡邊 [45] は、HRTF を共通極・零点モデルとしてモデル化したのち、そのパラメータを内挿する方法を提案した。これらの方法では、実際に測定する HRTF の割合と内挿される HRTF の推定精度とがトレードオフの関係にあることが課題となっている。

## 2.5.2 バイノーラル再生法

2.3 節でバイノーラル再生法を概説したが、この手続きにより得られるバイノーラル再生の聴感には、実音源からの再生を直接受聴することにより得られる聴感とはやや異なることが知られている。その原因として次のようなものが考えられる [32]。

まず最も原初的な原因として挙げられるのが、HRTF の個人性である。頭部や胴部、耳介の大きさや形状は個人により異なるため、HRTF も個人により異なっている。人は自身の HRTF と音の入射方向との関係を学習する。そのため、HATS や他者の HRTF を用いて生成されたバイノーラル信号を受聴するとき、提示される音像と知覚される音像の定位にずれが生じることになる。

次に挙げられるのは、音響無響室で測定された HRTF には再生音場での反射や響きの応答が含まれていないことである。反射音構造により得られる再生音場の空間情報は、音像の定位知覚、特に距離知覚における手掛かりの一つである。この情報の欠落は、受聴者に空間印象を錯覚させ、頭内定位や前後誤判定 (Front-Back Confusion) を引き起こすと考えられる。

受聴者の頭部運動が反映されないことも、原因の一つとして挙げられる。この説明としてしばしば用いられるのが混同の円錐 (Cone of Confusion) である [12,46]。頭部を球形モデルで近似したとき、両耳中心位置を頂点として両耳を通過する直線を軸とする円錐上の点に対し、両耳は常に一定の距離差となる。これは同円錐上に音源があるとき ITD が常に一定であることを意味しており、音像の定位知覚に曖昧さをもたらす。このようなとき、人は意識、無意識下での頭部運動により両耳と音源との相対位置関係を変化させ、これに伴う ITD や ILD、スペクトルパターンの変化を曖昧さ解消のための情報として収集していると考えられる。頭部の運動が許容されるとき音像の定位知覚が改善されるという実験結果は多く報告されており [47]、人が学習するその手掛かりも、自然な頭部運動や姿勢変化の中で獲得されると考えるのが自然である。

最後に挙げられるのは、再現される HRTF の不完全性である。マイクロフォンにより HRTF を測定した収音点とヘッドフォンやイヤフォンによりバイノーラル信号を提示する再生点が一致しない、測定系や再生系の特性が正確に補正されないなど、様々な要因により、測定された HRTF が再生時に正確に復元されないことがある。そのため、HRTF 測定時のマイクロフォンの据え付け法、イヤフォンや外耳道内特性の補正法などの研究が盛んに行われている [48, 49]。

これらの原因により引き起こされる定位知覚のずれや様々な誤判定は、よりよいバイノーラル再生を実現させようという動機付けとなっている。HRTF の測定や測定系、再生系の補正精度を高め、バイノーラル信号をできるだけ忠実に再現しようという物理的なアプローチ、妥当な聴こえを生じさせる手掛かりに注目し、それら手掛かりから聴覚系での認知に至るメカニズムを解明しようとする心理・神経生理学的なアプローチなど、バイノーラル再生法に係る諸検討は、現在も精力的に取り組まれている。

### 2.5.3 トランスオーラル再生法

トランスオーラル再生法においても、バイノーラル再生法と同様に、HRTF の個人性や不完全性による再生品質の低下がしばしば問題となる。HRTF の不一致は音像の定位感にずれを生じさせ、受聴者に不自然な聴感を与えることにより頭内定位や前後誤判定を引き起こす。加えて、これらの個人性や不完全性、受聴環境の変化に起因する再生音場の不確かさは、そのまま設計される制御器の不確かさとなる。この不確かさによりクロストークが効率的に抑制されなくなると、トランスオーラル再生自体が成立しなくなり、結果として音像が 2 次音源に定位して知覚されるようになる。

また、制御器のゲインはその不確かさや入力信号に混入する雑音成分を増幅する。増幅された不確かさや雑音成分は受聴者に歪みやアーティファクトとして知覚され、振幅の大きな周波数成分が含まれるような場合には、人の聴覚機構を損傷する危険もある。特に、処理過程に逆システムを含むトランスオーラル再生法の場合、再生音場の音響伝達関数のノッチが、逆システムに思わぬピークを形成することがある。そのため、それらのピークが大振幅とならないよう、ワーストケースを想定して制御器を設計する必要がある。

Nelson, 浜田らの研究グループは、制御器となる逆システムの設計に Tikhonov の正則化法を導入し、逆システムのゲインにペナルティを課すことによりその増幅を抑制する方法を提案した [50–57]。周波数伝達関数に基づく設計では、正則化法の基本式は

$$\mathbf{H}_{\text{reg}}(n) = \mathbf{G}_p^H(n) \{ \mathbf{G}_p(n) \mathbf{G}_p^H(n) + \beta(n) \mathbf{I} \}^{-1} \quad (2.21)$$

となる。ただし、 $\mathbf{G}_p(n) \in \mathbb{C}^{p \times m}$  は再生音場の周波数伝達関数行列を表し、式 (2.21) では  $p \leq m$  としている。また、 $\mathbf{H}_{\text{reg}}(n) \in \mathbb{C}^{m \times p}$  は正則化法により求められる逆システムを表す。

また,  $n$  は離散周波数ビンを表す.  $\beta(n)$  は正則化パラメータと呼ばれ, 逆システムのゲインに課されるペナルティの重みを表す. この値が大きくなるほど逆システムのゲインが抑制され, そのトレードオフとして精度は低下する. 彼らの研究グループは逆システムを良好条件で設計するための 2 次音源配置についても検討しており, Kirkeby, 浜田ら [52,53,55] はステレオダイポールを, Takeuchi ら [56,57] は OSD (Optimal Source Distribution) 原理に基づく OSD システムを提案した.

鹿野, 猿渡らの研究グループは, 逆システムを構成する伝達関数の線形独立性を高め, そのゲインを抑制する方法として, 打ち切り特異値分解 (Truncated Singular Value Decomposition: TSVD) 法の導入を提案した [8,58–61]. 周波数伝達関数に基づく設計では, まず, 再生音場の周波数伝達関数行列  $G_p(n)$  の特異値分解を行う.

$$G_p(n) = U(n)\Sigma(n)V^H(n) \quad (2.22)$$

ただし,  $U(n) \in \mathbb{C}^{p \times p}$ ,  $V(n) \in \mathbb{C}^{m \times m}$  はユニタリ行列である.  $\Sigma(n) \in \mathbb{R}^{p \times m}$  は  $G_p(n)$  の特異値  $\sigma_i(n)$  ( $i = 1, 2, \dots$ ) を要素として持つ行列である. 簡単のため  $p \leq m$  とし,  $\Sigma(n)$  上に特異値は値の大きいものから順に並んでいるとすると,

$$\Sigma(n) = \begin{bmatrix} \Sigma_p(n) & \mathbf{0}_{p \times (m-p)} \end{bmatrix} \quad (2.23)$$

となる. ただし,

$$\Sigma_p(n) = \text{diag}[\sigma_1(n), \dots, \sigma_p(n)], \quad \sigma_1(n) > \sigma_2(n) > \dots > \sigma_p(n) \quad (2.24)$$

である. 次に,  $\Sigma_p(n)$  の逆数を対角成分に持つ対角行列  $\Gamma(n) \in \mathbb{R}^{p \times p}$  を計算する. このとき, 小さな特異値 (逆システムにおいては大きな特異値) を打ち切り 0 に置換する.

$$\Gamma(n) = \text{diag}[\gamma_1(n), \dots, \gamma_p(n)] \quad (2.25)$$

ただし,

$$\gamma_i(n) = \begin{cases} \frac{1}{\sigma_i(n)}, & \sigma_i(n) > \text{threshold} \\ 0, & \text{otherwise} \end{cases} \quad (2.26)$$

である. 式 (2.25) を用いて, 逆システム  $H_{\text{tsvd}}(n) \in \mathbb{C}^{m \times p}$  は

$$H_{\text{tsvd}}(n) = V(n) \begin{bmatrix} \Gamma(n) \\ \mathbf{0}_{(m-p) \times p} \end{bmatrix} U^H(n) \quad (2.27)$$

により与えられる. 得られた逆システム  $H_{\text{tsvd}}(n)$  は打ち切り数に応じてそのゲインが抑制されると共に, フロベニウスノルムの意味で元の逆システムの最良近似となる. なお, 打ち切り数の決定については, あらかじめその数を決めておく方法, 特異値を打ち切りの閾値として用いる方法のほか, 最大特異値と最小特異値の比として計算される条件数を打ち

切りの閾値として用いる方法 [62]，再生音場のノイズ拡大と関連付けて打ち切り数を決定する方法 [8] などが検討されている。

正則化法および TSVD 法はいずれも，逆システムの精度をある程度犠牲にして良条件での求解を可能とする，緩和処理法と考えることができる。比較的少ない計算量で実行可能であることから，トランスオーラル再生法に限らず様々な音場再現・再生法で応用されている。例えば李ら [63] は，立体音響再生における多チャンネル逆システムの設計に正則化法を用い，逆システムの再現精度を評価指標として正則化パラメータを決定する方法を提案した。田中ら [64] は，雑音成分の分散を評価指標として，同じく正則化パラメータを決定する方法を提案した。また，浅井ら [62] は境界音場制御における再生音場の特性のキャンセルに，Kurabayashi ら [65] は頭部運動を許容する動的トランスオーラル再生法の設計に，それぞれ TSVD 法を用いている。一方，近似を伴わない逆システムの設計法としては，Miyoshi ら [66] が室内音響の逆フィルタリング法として提案する MINT (Multiple-Input/Output Inverse Theorem) 法の応用が考えられる。

そのほか，近年ではロバスト制御理論の応用も検討され始めており，Samejima ら [11]，佐藤ら [67] はともに， $\mathcal{H}_\infty$  制御による逆システムの設計法を提案している。

これら従来法はいずれも，非最小位相系である再生音場の逆システムを制御器としている。そのため，因果性を満たすために遅延を伴う制御となること，緩和処理によりゲインを抑制する際に制御器の精度低下が避けられないことが課題となっている。また，逆システムのゲインを陽に管理するための手続きが含まれていないため，得られる解が必ずしも最適解となるとは限らない点も検討課題となっている。

以上整理した課題を背景として，本論文では，HRTF の測定時間を短縮する多方向同時推定法と，トランスオーラル再生制御器のゲインを抑制しつつ，制御に係る遅延を最小限に抑えたトランスオーラル再生法を提案することを目的とする。



## 第 3 章

# システム同定理論に基づく頭部伝達関数の多方向同時推定

頭部伝達関数 (Head-Related Transfer Function: HRTF) の測定では、測定方向数に比例して増加する測定時間が課題となる。人の特性を実測する場合、被測定者は終始姿勢を正した状態を維持する必要があるため、長時間の測定は身体的負担となることも考えられる。

本章では、多方向の頭部インパルス応答 (Head-Related Impulse Response: HRIR) を MISO FIR モデルとみなし、システム同定理論に基づいて同時に推定する方法を提案する。MISO FIR モデルの推定にあたって、その同定入力の一つの M 系列から作成する方法が提案されている [68]。本章では、HRIR の推定に適用するために同方法を修正した、同定入力の作成法も提案する [69]。

### 3.1 予測誤差法による頭部伝達関数の同時推定

Fig. 3.1 に提案する同時推定法の概略図を示す。被測定者の周囲、HRTF の測定方向に音源となるスピーカを配置する。このとき、各測定方向から左右耳までの HRIR の集合は、音源を入力端、左右耳の外耳道入口を出力端とする多入力多出力 (Multi Input Multi Output: MIMO) FIR システムと考えることができる。更に、左右耳の HRIR は独立したシステムとなることから、左耳と右耳のそれぞれの外耳道入口を出力端とする二つの MISO FIR システムと考えることができる。図中、 $y(k)$  は離散時間の時刻  $k$  における左耳での出力、 $u_i(k)$  ( $i = 1, 2, \dots, m$ ) は  $i$  番目の入力を表す。 $m$  は音源の数、すなわちシステムの入力数である。提案法では、この音源より測定信号を同時に再生し、被測定者の両耳外耳道入口でその応答を收音する。この測定信号と收音された応答を同定入出力  $u_i(k)$  ( $i = 1, 2, \dots, m$ ) および  $y(k)$  として、左右耳の HRIR を予測誤差法により推定する。

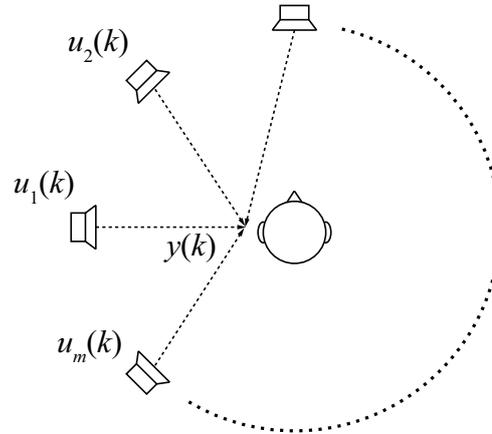


Fig. 3.1: Schematic diagram illustrating simultaneous estimation of HRTFs

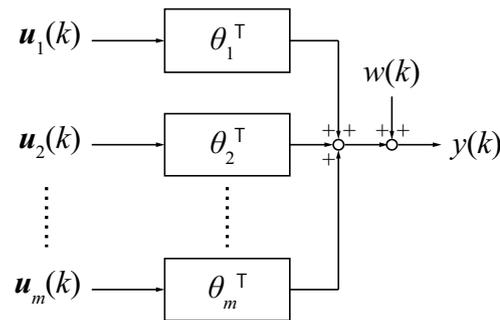


Fig. 3.2: Block diagram illustrating simultaneous estimation of HRTFs

$m$  入力 1 出力の  $n$  次 MISO FIR モデルは

$$y(k) = \theta_1^T \mathbf{u}_1(k) + \theta_2^T \mathbf{u}_2(k) + \cdots + \theta_m^T \mathbf{u}_m(k) + w(k) \quad (3.1)$$

と表される．ただし， $\mathbf{u}_i(k)$  は  $i$  番目の入力  $u_i(k)$  から構成される入力ベクトル， $\theta_i$  は  $i$  番目の FIR モデルのパラメータベクトル，すなわち  $i$  番目の方向の HRIR を表し，それぞれ

$$\mathbf{u}_i(k) = \begin{bmatrix} u_i(k-1) \\ u_i(k-2) \\ \vdots \\ u_i(k-n) \end{bmatrix}, \quad \theta_i = \begin{bmatrix} \theta_{i1} \\ \theta_{i2} \\ \vdots \\ \theta_{in} \end{bmatrix} \quad (3.2)$$

である．また， $w(k)$  は出力に混入する正規性白色雑音を表す．式 (3.1) をブロック線図で表したものが Fig. 3.2 である．入出力信号長を  $N$  として式 (3.1) を時刻  $k = 1, \dots, N$  まで

並べると，

$$\mathbf{y}_N = \mathbf{U}_N \boldsymbol{\theta} + \mathbf{w}_N \quad (3.3)$$

と書き表すことができる．ただし，

$$\mathbf{y}_N = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}, \quad \mathbf{U}_N = \begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \\ \vdots \\ \mathbf{U}_m \end{bmatrix}^T, \quad \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_m \end{bmatrix}, \quad \mathbf{w}_N = \begin{bmatrix} w(1) \\ w(2) \\ \vdots \\ w(N) \end{bmatrix} \quad (3.4)$$

であり，

$$\mathbf{U}_i = \begin{bmatrix} u_i(0) & u_i(1) & \cdots & u_i(N-1) \\ u_i(-1) & u_i(0) & \cdots & u_i(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ u_i(1-n) & u_i(2-n) & \cdots & u_i(N-n) \end{bmatrix} \quad (3.5)$$

である．式 (3.3) は，時刻  $k-1$  までの過去の入出力と現時刻  $k$  の出力との関係を表している．この関係式に基づく出力  $\mathbf{y}_N$  の 1 段先予測値  $\hat{\mathbf{y}}_N$  は

$$\hat{\mathbf{y}}_N = \mathbf{U}_N \boldsymbol{\theta} \quad (3.6)$$

により与えられる．

予測誤差法とは，この予測値に基づく予測誤差から構成される評価関数を最小化する推定値を計算する，パラメータ推定法の総称である [71]．予測誤差の大きさの測度として 2 次関数を用いると最小 2 乗法となる．このとき，評価関数  $J_N$  は

$$J_N(\boldsymbol{\theta}) = \frac{1}{N} \|\mathbf{y}_N - \mathbf{U}_N \boldsymbol{\theta}\|_2^2 \quad (3.7)$$

となる．これを最小化する最小 2 乗推定値  $\hat{\boldsymbol{\theta}}$  は

$$\hat{\boldsymbol{\theta}} = \mathbf{R}_N^{-1} \mathbf{f}_N \quad (3.8)$$

により計算される．ただし，

$$\mathbf{R}_N = \frac{1}{N} \mathbf{U}_N^T \mathbf{U}_N, \quad \mathbf{f}_N = \frac{1}{N} \mathbf{U}_N^T \mathbf{y}_N \quad (3.9)$$

とおいた．

## 3.2 同定入力作成

式 (3.8) が最小 2 乗推定値として解を持つためには，同定入力  $u_i(k)$  ( $k = 1, \dots, N$ ) により構成される行列  $\mathbf{R}_N$  が正定値行列である必要がある．この条件を満たす入力は，次の補題 3.1 に基づいて作成することができる [68, 72]．

## 補題 3.1

$m$  入力 1 出力  $n$  次 FIR モデルは、次の条件を満たすとき、周期  $T$  の  $M$  系列を  $l$  サンプルずつ巡回シフトさせて作成した  $m$  個の入力を用いて同定することができる。

$$l > n \quad (3.10)$$

$$T > n + (m - 1)l \quad (3.11)$$

補題 3.1 は、FIR モデルの次数が既知であること、この FIR モデルの次数を基準として決められる巡回シフト幅内に入力に対する応答が 0 に収束することを仮定している。HRIR の場合、その長さは一般に、48 kHz のサンプリング周波数で 128 ~ 512 サンプルとされている。したがって、FIR モデルの次数は既知としてよく、測定を残響のない音響無響室で実施し、かつ測定治具などでの反射を抑制することができるのであればこの仮定が成立する。しかし、この仮定が成立しないとき、すなわち巡回シフト幅より遅れて反射や残響が混入するときには、隣接する入力の応答が相関を持つようになるため良条件での求解が困難となり、FIR モデルの推定精度が大幅に低下する。そこで補題 3.1 に、入力の巡回シフト幅  $l$  が FIR モデルの次数  $n$  より大きく、かつ応答の収束時間（サンプル数） $n_c$  よりも大きいという制約を追加する。この追加により、次に示す定理 3.1 が導かれる。

## 定理 3.1

$m$  入力 1 出力  $n$  次 FIR モデルは、次の条件を満たすとき、周期  $T$  の  $M$  系列を  $l$  サンプルずつ巡回シフトさせて作成した  $m$  個の入力を用いて同定することができる。ただし、 $n_c$  は応答の収束時間（サンプル数）である。

$$l \geq \max\{n_c, n\} \quad (3.12)$$

$$T \geq \max\{n_c, n\} + (m - 1)l \quad (3.13)$$

以下に入力の具体的な作成手続きを示す。なお、推定する FIR モデルの次数  $n$  は既知とする。

1. FIR モデルの次数  $n$  および応答の収束時間  $n_c$  より大きい値として、巡回シフト幅  $l$  を定める。ただし、 $l$  を大きくすると、それに比例して  $M$  系列の周期  $T$  も長くする必要があるので留意する。
2. 次式に示す原始多項式より、シフトレジスタ数  $n_s$  の  $M$  系列を作成する。

$$m_s(k) = m_s(k - 1) \oplus m_s(k - n_s) \quad (3.14)$$

ただし、 $\oplus$  は排他的論理和を表す。また、 $T = 2^{n_s} - 1$  は、 $l$  サンプルの巡回シフトを入力数  $m - 1$  回繰り返してもなお一巡しないよう十分に大きくする。予測誤差法の計算過程で必要となる計算時間および計算機上の作業領域は、 $T$  の大きさに応じ

て増加する．したがって，測定時間と後処理時間，求められる推定精度，後処理に使用する計算機のスペックを総合的に判断して  $l$  と  $n_s$  を定める必要がある．

3. 作成した  $M$  系列を 1 番目の入力

$$u_1(k) = m_s(k) \quad (3.15)$$

とする．

4. 1 番目の入力  $u_1(k)$  を時間軸上で  $l$  サンプルだけ負の方向に巡回シフトしたものを 2 番目の入力

$$u_2(k) = \begin{cases} m_s(k+l), & 1 \leq k \leq T-l \\ m_s(k+l-T), & T-l < k \leq T \end{cases} \quad (3.16)$$

とする．

5. 以下同様に， $i$  番目 ( $i = 2, 3, \dots, m$ ) の入力  $u_i(k)$  は， $i-1$  番目の入力  $u_{i-1}(k)$  を時間軸上で  $l$  サンプルだけ負の方向に巡回シフトしたものとする．

$$u_i(k) = \begin{cases} m_s(k+(i-1)l), & 1 \leq k \leq T-(i-1)l \\ m_s(k+(i-1)l-T), & T-(i-1)l < k \leq T \end{cases} \quad (3.17)$$

この手続きに従って作成した入力を用いることより，隣接する応答への反射や残響の混入が回避され，FIR モデルの推定精度が改善される．

### 3.3 実験 1: 頭部伝達関数の多方向同時推定実験

提案する HRTF の同時推定法の有効性を検証するため，1 方向ずつの測定との比較実験を実施した．本節ではまず，補題 3.1 に基づいて同定入力を作成し，これを測定信号として同時推定に用いた場合の結果を示す．リファレンスとなる 1 方向ずつの測定には OATSP を測定信号として用いた．また，リファレンスと比較しての誤差の評価指標にはスペクトル歪み (Spectral Distortion: SD) を用いた<sup>\*1</sup>．

#### 3.3.1 実験条件

実験は，NHK 放送技術研究所の防音室で実施した (Fig. 3.3)．防音室の残響時間は 500 Hz で 0.18 s である．Fig. 3.4 に音源となるスピーカの配置を示す．被測定者の両耳中心位置を含む水平面を仰角  $\varphi = 0^\circ$  として， $\varphi = 0^\circ$  の層には 24 台， $\varphi = 30^\circ, -30^\circ$  の層には各 12 台， $\varphi = 60^\circ$  の層には 8 台，そして天頂となる  $\varphi = 90^\circ$  の層には 1 台の計 57 台

<sup>\*1</sup> 文献によりスペクトル距離 (Spectral Distance: SD) と記述されることもある．定義は同一であり，その略記も同じ SD となる．



Fig. 3.3: Measurement of impulse responses for simultaneous estimation in a sound proof room of the NHK Science and Technology Research Laboratories

Table 3.1: Experimental conditions in experiment 1

Site	Soundproof room at NHK STRL
Reverberation time (s)	0.18 at 500 Hz
Number of estimation directions	57 directions
Number of subjects	20 human subjects and HATS (KEMAR Type 45BA)

のスピーカを，球上 5 層に渡って等間隔で配置した．なお，被測定者の正面方向を方位角  $\theta = 0^\circ$  とし，天頂より見て反時計回りに正とした．各スピーカから被測定者の両耳中心位置までの距離は 1.8 m である．被測定者の周囲，測定方向に多数配置したスピーカとその治具は反射物となるため，吸音材を巻いてその影響を低減した．

実験条件を Table 3.1 にまとめる．被測定者は HATS と成人被測定者 20 名とした．HATS での測定にはイヤースミュレータと外耳道入口に内蔵されるコンデンサマイクروفオンを使用し，人の実測では第 2 章で説明したバイノーラルマイクروفオンを使用した．

測定信号を Table 3.2 にまとめる．同時推定に用いる同定入力は，シフトレジスタ数  $n_s = 15$  の M 系列より作成した．このとき信号長  $N$  (周期  $T$ ) は  $2^{15} - 1$  サンプルとなる．サンプリング周波数は 48 kHz，量子化ビット数は 24 ビットとした．FIR モデルの次数

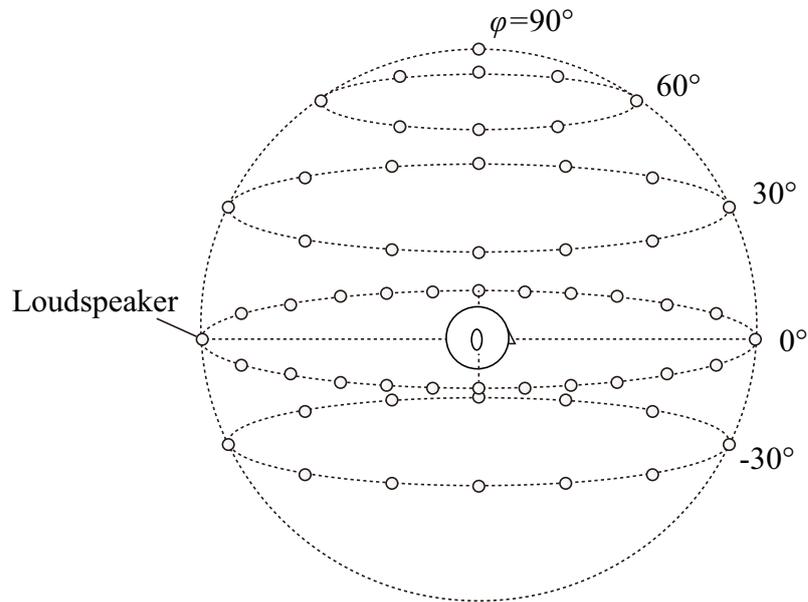


Fig. 3.4: Loudspeaker arrangement

Table 3.2: Input signals in experiment 1

	MLS	OATSP
Sampling frequency (kHz)	48	48
Bit depth (bits)	24	24
Signal length (sample)	$2^{15} - 1$	$2^{16}$
Cyclic shift length (sample)	520	-
Model order	128	-
Sound pressure level (dB, A-weighted)	67	65

$n$  は主応答が十分に収束する 128 次とし，巡回シフト幅  $l$  はこれより長い 520 サンプルとした．また，リファレンスとする 1 方向ずつの測定で用いる OATSP の信号長は M 系列より長い  $2^{16}$  サンプルとした．サンプリング周波数，量子化ビット数は M 系列と同じ 48 kHz，24 ビットとした．

同時推定では，測定信号となる同定入力をスピーカより同時に再生し，その応答を左右耳位置のマイクロフォンで收音した．一方，1 方向ずつの測定では，測定信号を各スピーカから順番に再生し，応答をその都度收音した．ただし，いずれの測定でも，スピーカとマイクロフォンの特性により，応答の帯域を 20 Hz から 20 kHz に制限した．実験音圧レベルは，同時推定ではピークを 67 dB (A-weighted)，1 方向ずつの測定ではピークを

65 dB (A-weighted) とし、いずれも被測定者がいない状態での両耳中心位置において調整した。通常の測定ではスピーカが歪まない範囲で測定系のゲインが最大となるよう調整するが、測定が長時間に及ぶため、被測定者の負担を考慮してやや低めの音圧とした。

### 3.3.2 実験結果

推定例として、水平面上  $(\theta, \varphi) = (0^\circ, 0^\circ), (90^\circ, 0^\circ), (180^\circ, 0^\circ), (270^\circ, 0^\circ)$  の音源から HATS への HRIR の周波数特性を Fig. 3.5 に、正中面上  $(\theta, \varphi) = (0^\circ, -30^\circ), (0^\circ, 30^\circ), (0^\circ, 60^\circ), (0^\circ, 90^\circ)$  の音源から HATS への HRIR の周波数特性を Fig. 3.6 に示す。図中、赤実線は OATSP を用いて 1 方向ずつ測定した HRIR の周波数特性であり、青実線は M 系列を用いた同時推定により得られた HRIR の周波数特性である。なお、全 57 方向の HRIR の振幅のピークが  $-2$  dB (Full Scale) となるよう正規化した。Fig. 3.5 は推定精度に方位依存性があること示しており、(c) (d) (g) (h) を見ると、音源と同側耳の HRIR は精度よく推定されているが、頭部の陰となる対側耳の HRIR は 1 方向ずつ測定した HRIR から誤差が生じている。一方、Fig. 3.6 では、推定精度に仰角依存性はそれほど見られなかった。

この誤差を定量的に評価するため、次式により定義される SD を物理評価指標として導入する。

$$SD = \sqrt{\frac{1}{N} \sum_{n=1}^N \left( 20 \log_{10} \frac{|H_{OATSP}(n)|}{|H_M(n)|} \right)^2} \quad (\text{dB}) \quad (3.18)$$

ただし、 $n$  は離散周波数ピンを、 $N$  はその数を表す。SD は二つの振幅スペクトルの差異を対数軸上で評価し、全周波数帯域において平均化したものである。その値が小さいほど誤差が小さいことを示し、人の聴覚特性とよく整合する尺度として、音声情報処理などの分野でよく用いられる。

まず、HATS の HRIR について、水平面各方向の SD を Fig. 3.7 に、正中面各方向の SD を Fig. 3.8 に示す。Fig. 3.7 は、同側方向の HRIR では小さく抑えられている誤差が、対側方向に移るに従い徐々に大きくなることを示している。これは、主応答の収束後に到来する残響、反射の影響により、振幅の小さい HRIR から SN 比が低下したためと考えられる。左右耳で SD の推移が非対称となっているが、これは HATS (KEMAR Type 45BA) およびイヤーマイクの形状が、人の身体と同様に左右非対称になっているためである。一方、Fig. 3.8 では、大きな仰角依存性は見られなかった。なお、左耳での平均 SD は 2.6 dB、右耳での平均 SD は 2.7 dB であった。

次に、20 名の被測定者の HRIR について、左右耳それぞれ 57 方向、計 114 方向の平均 SD の分布を Fig. 3.9 に示す。各被測定者の平均 SD は 2.5 ~ 3.5 dB の範囲にあり、中央値は 2.9 dB であった。

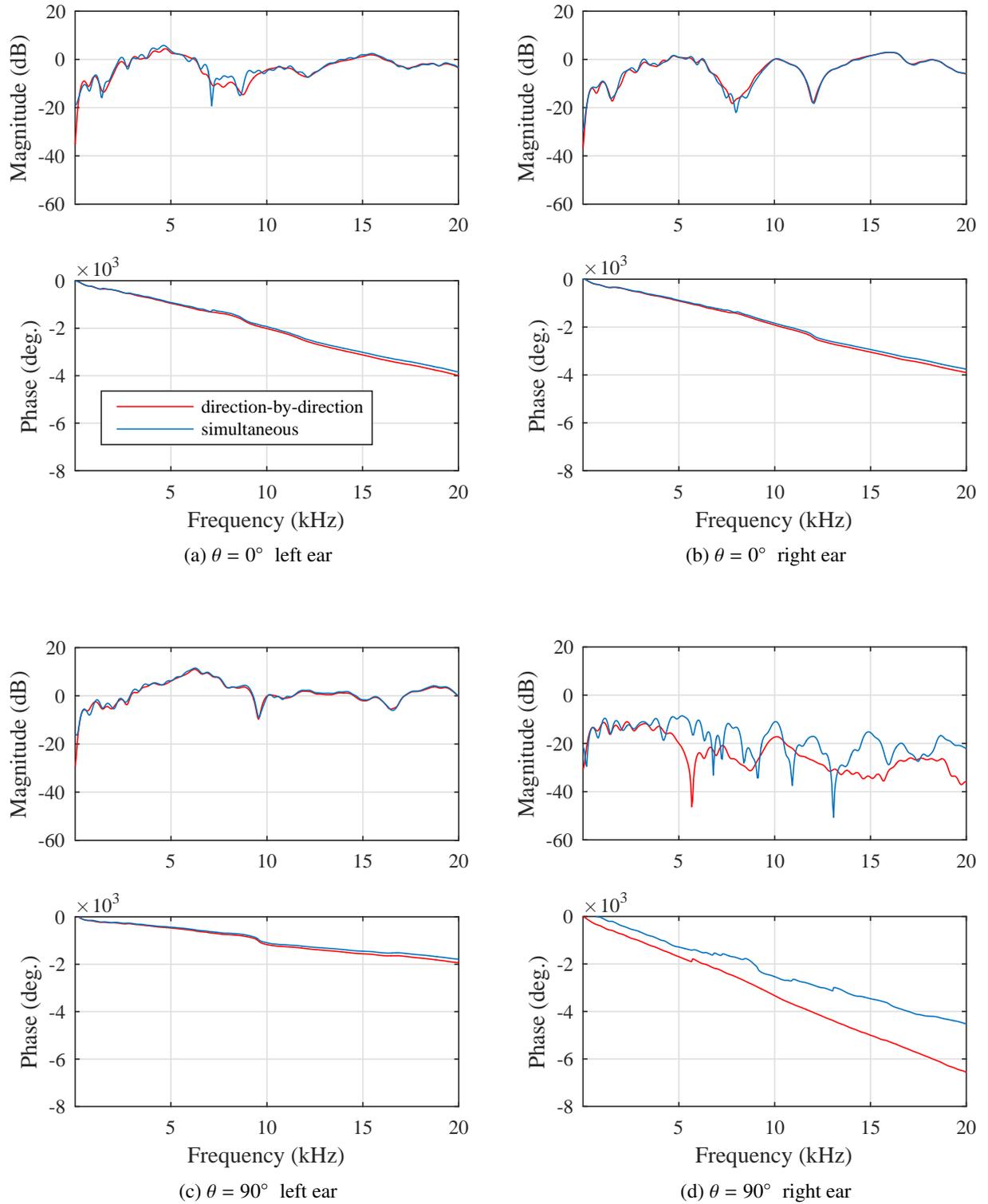


Fig. 3.5: Frequency characteristics of estimated HRIRs (horizontal plane, red solid line: direction-by-direction measurement, blue solid line: simultaneous estimation)

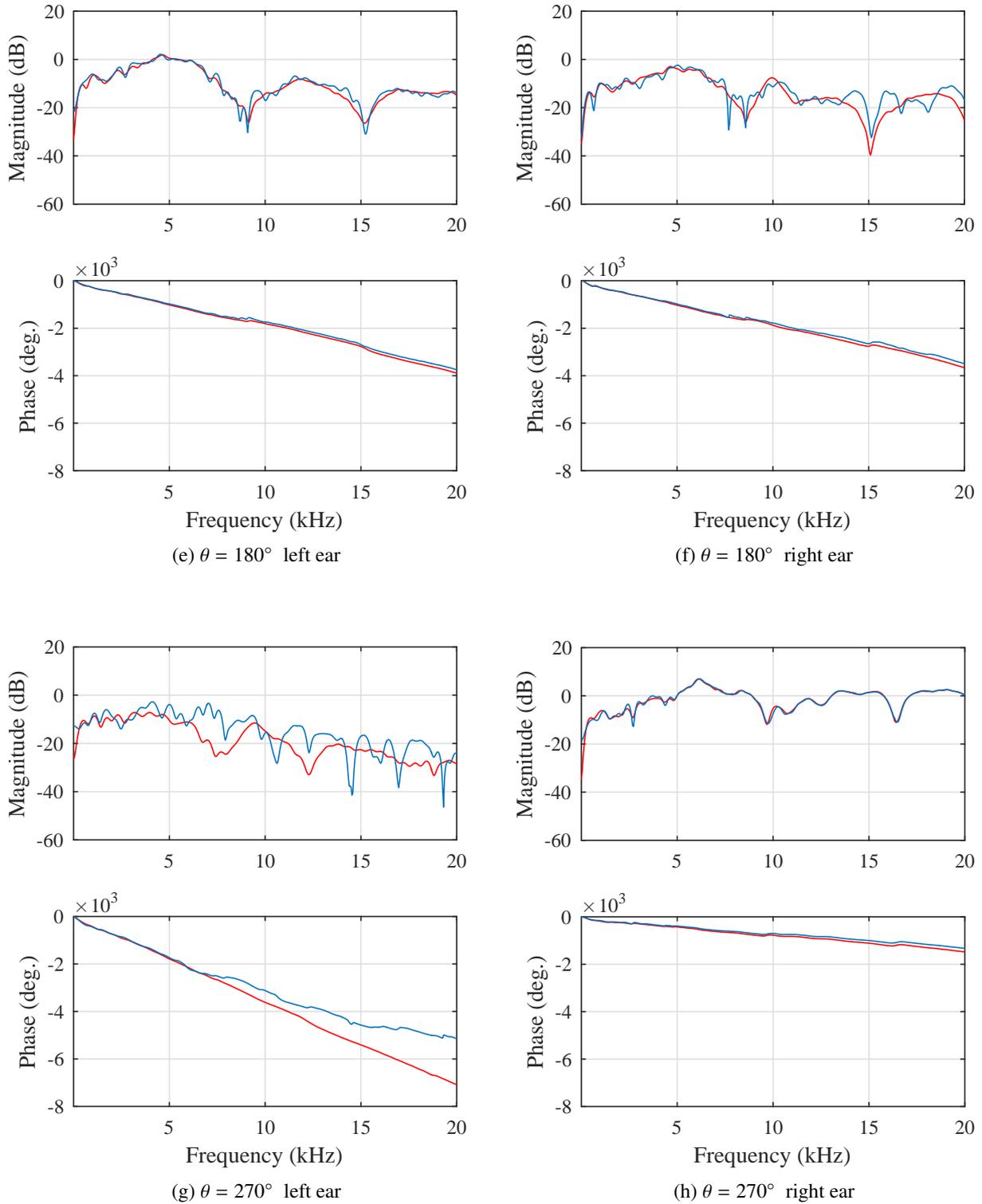


Fig. 3.5: Frequency characteristics of estimated HRIRs (horizontal plane, red solid line: direction-by-direction measurement, blue solid line: simultaneous estimation)

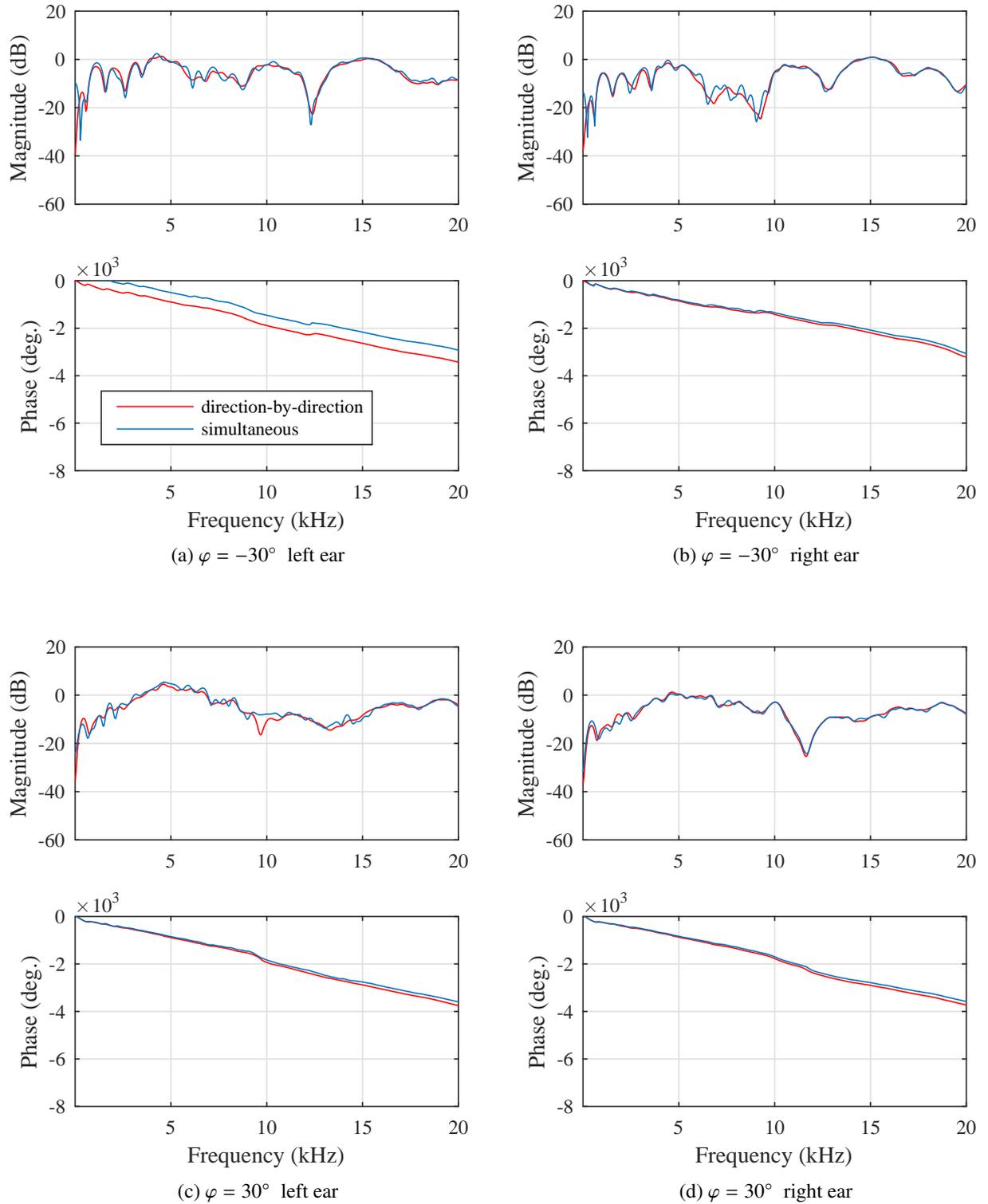


Fig. 3.6: Frequency characteristics of estimated HRIRs (median plane, red solid line: direction-by-direction measurement, blue solid line: simultaneous estimation)

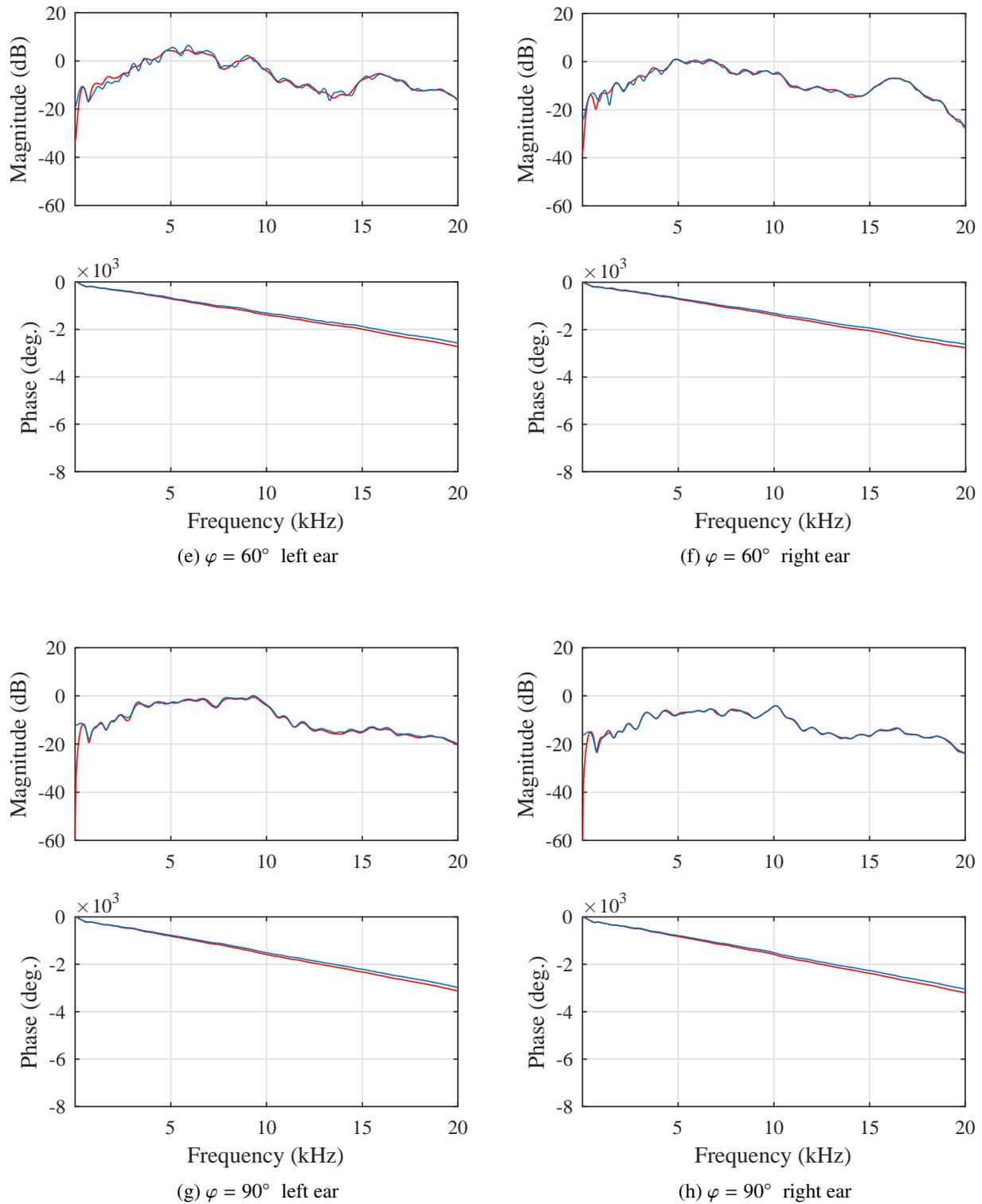


Fig. 3.6: Frequency characteristics of estimated HRTFs (median plane, red solid line: direction-by-direction measurement, blue solid line: simultaneous estimation)

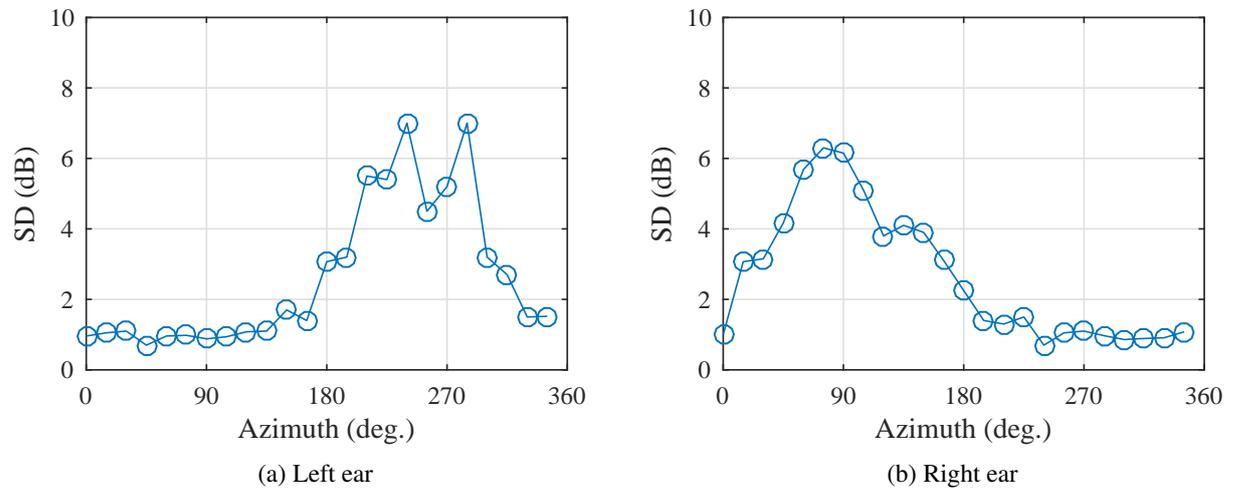


Fig. 3.7: SDs of estimated HRTFs (horizontal plane)

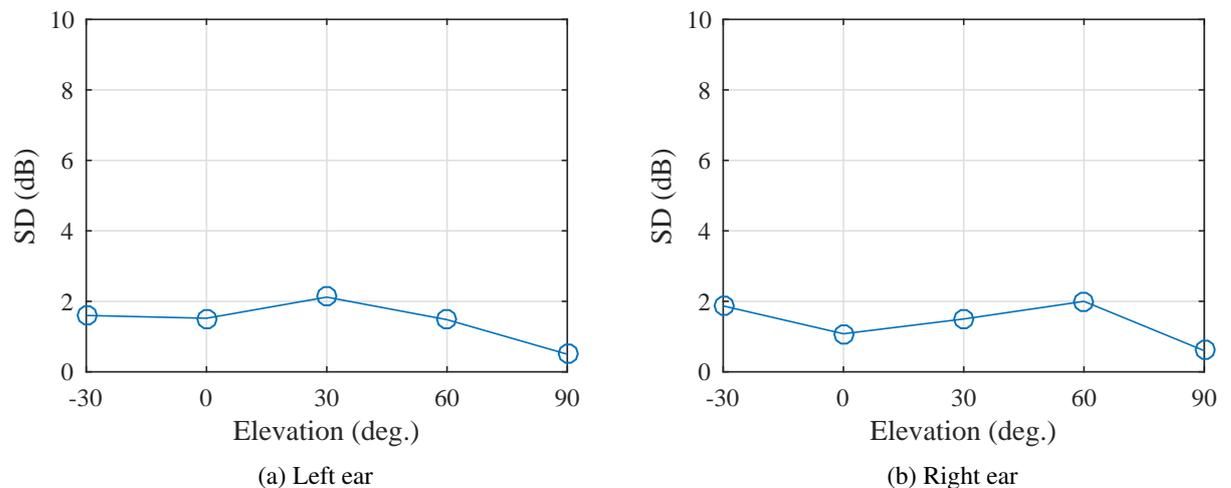


Fig. 3.8: SDs of estimated HRTFs (median plane)

人の聴覚における空間分解能は方向によって異なっており、HRIRの畳み込まれた音源の歪みの検知限も音源の方向や種類、運動の有無などの諸条件により異なっている。また、歪みを許容する程度も受聴者の訓練度合いにより異なることから、SDの許容限を一概に定めることは困難である。例えば、HRIRを畳み込んだ音源の移動感を表現するためには、SDは6.0 dB程度に抑えられる必要があるとされている。また、音像の定位知覚される方向を正しく表現するためには、これより低く3 dB以下とされている[73]。自身のHRTFを基準として許容される歪みを調べると、SDは更に低くなり、音源の種類への依

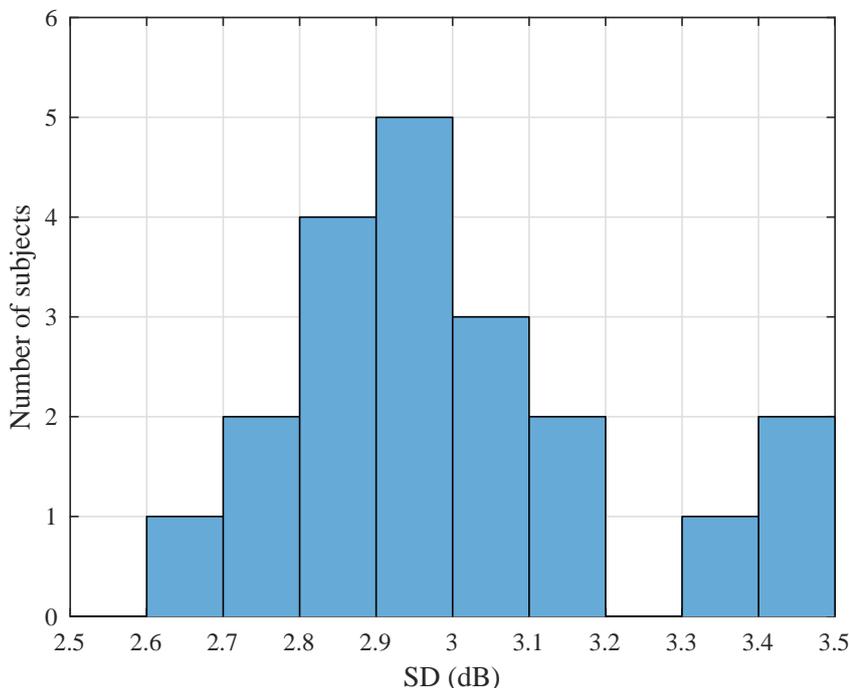


Fig. 3.9: Distribution of average SDs of 20 subjects

存が高くなる．例えば，音源として楽音を用いた音像定位実験では，同側耳で 2 ~ 3 dB 程度，対側耳で 5 ~ 7 dB が SD の許容限となる．音源として白色雑音を用いた場合では，おおよそ 2 dB 以下となる [74]．これらの知見に鑑みるに，音像の定位感について「その差は分かるが気にならない」許容限の範囲にある推定精度である．

### 3.4 実験 2: 提案する同定入力を用いた多方向同時推定実験

実験 1 では，補題 3.1 に基づいて巡回シフト幅を 520 サンプルとして同定入力を作成した．これは，サンプリング周波数を 48 kHz としたときの HRIR のサンプル数として，一般的に 128 ~ 512 サンプルが選ばれることを根拠としている．しかし，防音室の残響時間は 500 Hz で 0.18 s であることから，主応答の収束後に到来する残響，反射の影響により，特に頭部の陰となる対側耳において HRIR の推定精度が低下した．

そこで本節では，定理 3.1 に基づいて巡回シフト幅をこの残響時間より大きく設定して同定入力を作成し，これを測定信号として同時推定に用いた場合の結果を示す．リファレンスとなる 1 方向ずつの測定には実験 1 と同様に OATSP を測定信号として用いた．併せて補題 3.1 に基づいて同定入力を作成し，これを測定信号とした同時推定と比較することにより，定理 3.1 の妥当性と有効性を検証した．

Table 3.3: Experimental conditions in experiment 2

Site	Soundproof room at NHK STRL
Reverberation time (s)	0.18 at 500 Hz
Number of estimation directions	24 directions on the horizontal plane
Subjects	HATS (KEMAR Type 45BA)

Table 3.4: Input signals in experiment 2

	MLS (8,640)	MLS (520)	OATSP
Sampling frequency (kHz)	48	48	48
Bit depth (bits)	24	24	24
Signal length (sample)	$2^{18} - 1$	$2^{18} - 1$	$2^{18}$
Cyclic shift length (sample)	8,640	520	-
Model order	128	128	-
Sound pressure level (dB, A-weighted)	65	65	65

### 3.4.1 実験条件

実験は、実験 1 と同じ NHK 放送技術研究所の防音室で実施した。音源となるスピーカは仰角  $\varphi = 0^\circ$ 、水平面上の 24 台を使用した。

実験条件を Table 3.3 にまとめる。被測定者は HATS とした。

測定信号を Table 3.4 にまとめる。同時推定に用いる測定信号は、巡回シフト幅を大きく設定する必要から、シフトレジスタ数  $n_s = 18$  の M 系列より作成した。このとき信号長  $N$  (周期  $T$ ) は  $2^{18} - 1$  サンプルとなる。サンプリング周波数、量子化ビット数は実験 1 と同じ 48 kHz, 24 ビットとした。FIR モデルの次数  $n$  は 128 次とし、巡回シフト幅  $l$  は式 (3.12) に従い防音室の残響時間と等しい 8,640 サンプルとした。補題 3.1 に基づく推定では、 $l$  は実験 1 と同じ 520 サンプルとした。また、リファレンスとする OATSP の信号長、サンプリング周波数、量子化ビット数は M 系列と同じ  $2^{18}$  サンプル、48 kHz, 24 ビットとした。

実験音圧レベルはピークを 65 dB (A-weighted) とし、HATS を置かない状態での両耳中心位置において調整した。

### 3.4.2 実験結果

推定例として、水平面上  $(\theta, \varphi) = (0^\circ, 0^\circ), (90^\circ, 0^\circ), (180^\circ, 0^\circ), (270^\circ, 0^\circ)$  の音源から HATS への HRIR の周波数特性を Fig. 3.10 に示す。図中、赤破線は OATSP を用いて 1 方向ずつ測定した HRIR の周波数特性である。赤実線は巡回シフト幅 520 サンプルの M 系列を用いた同時推定により得られた HRIR の周波数特性であり、青実線は巡回シフト幅 8,640 サンプルの M 系列を用いた同時推定により得られた HRIR の周波数特性である。なお、全 24 方向の HRIR の振幅のピークが  $-2$  dB (Full Scale) となるよう正規化した。

Fig. 3.10 は、いずれの方向においても HRIR が精度よく推定できていることを示している。特に Fig. 3.10 (d) (g) は、巡回シフト幅の修正により、頭部の陰となる対側耳においても同側耳に近い精度まで改善されたことを示している。このことから、定理 3.1 の有効性が証明されたと考える。なお、位相特性に見られるずれは、プロット時の位相アンラップによる誤差の累積が原因である。

各方向の SD を Fig. 3.11 に示す。図中、赤実線は巡回シフト幅 520 サンプルの M 系列を用いた同時推定の結果を表し、青実線は巡回シフト幅 8,640 サンプルの M 系列を用いた同時推定の結果を表す。巡回シフト幅の修正により、左耳での平均 SD は 2.90 dB から 0.35 dB に、右耳での平均 SD は 2.55 dB から 0.33 dB にそれぞれ低減された。これは、音像の定位感に加え、音色についても「その差が分かるが気にならない」許容限の範囲にある推定精度である [74]。

## 3.5 実験の考察

実験により、定理 3.1 に基づいて測定信号を作成すれば、音像の定位感、音色ともに十分許容される精度で HRTF を推定可能であることが示された。文献 [74] の実験に従えば、一対比較法や MUSHRA (Multiple Stimuli with Hidden Reference and Anchor) 法などによりリファレンスと比較試聴してはじめてその差異が知覚できる程度の誤差である。したがって、バイノーラル再生法、トランスオーラル再生法で用いるには十分な推定精度であると考えられる。

提案法の測定時間は、測定方向数によらず測定信号長に応答の収束時間を加えた時間となる。これに対して、1 方向ずつの測定では、測定信号長に応答の収束時間を加えた時間をインターバルとして、これを測定方向数分だけ繰り返す必要がある。したがって、その測定時間は、提案法によりおよそ測定方向数分の 1 まで短縮されることになる。巡回シフト幅により測定信号長は伸縮するが、全体の測定時間に対してその影響は僅かである。例えば実験 1 では、OATSP の信号長は約 1.36 s ( $\approx 2^{16}/48,000$ ) であり、インター

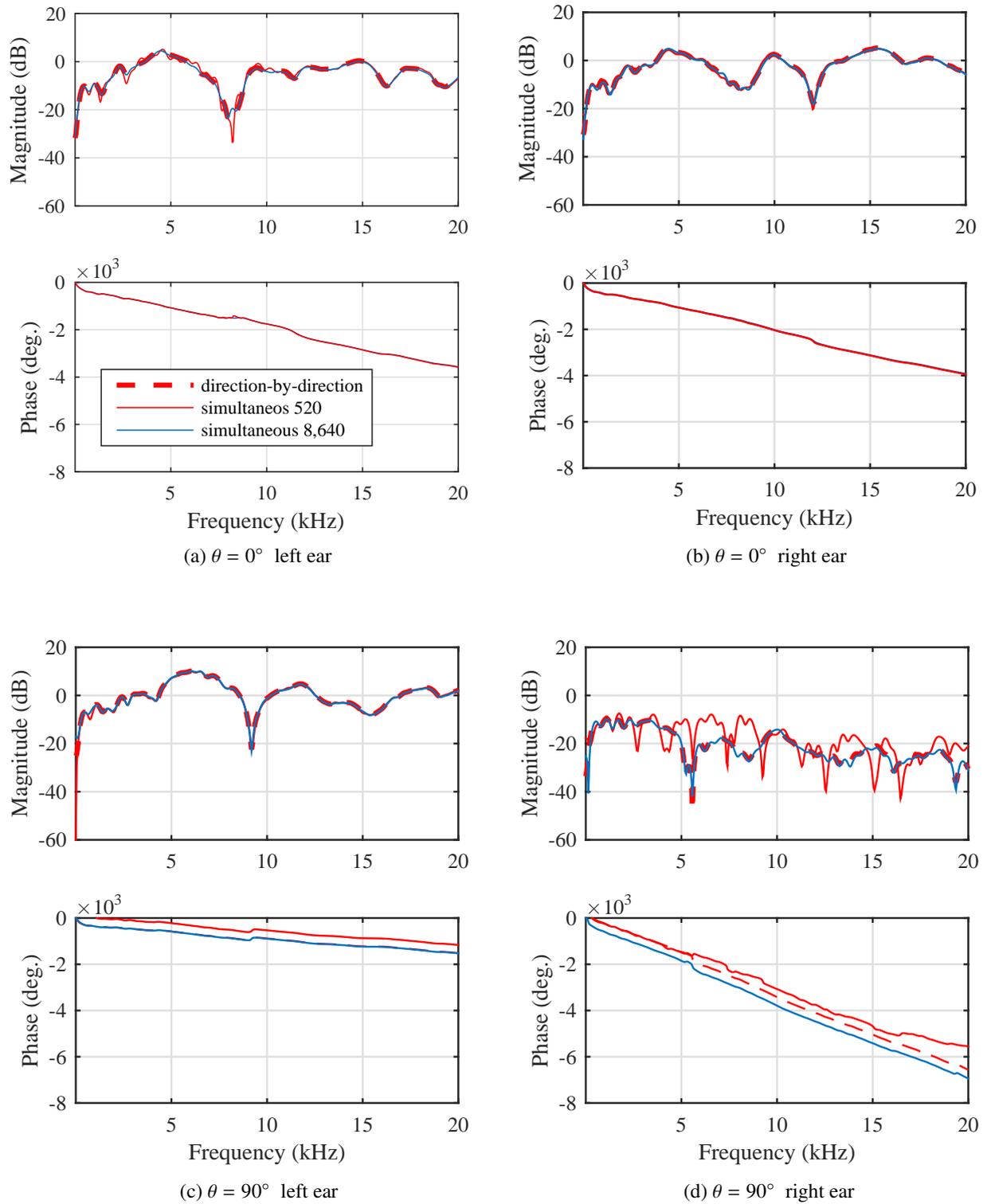


Fig. 3.10: Frequency characteristics of estimated HRIRs (red dashed line: direction-by-direction measurement, red solid line: simultaneous estimation with 520 sample shift, blue line: simultaneous estimation with 8,640 sample shift)

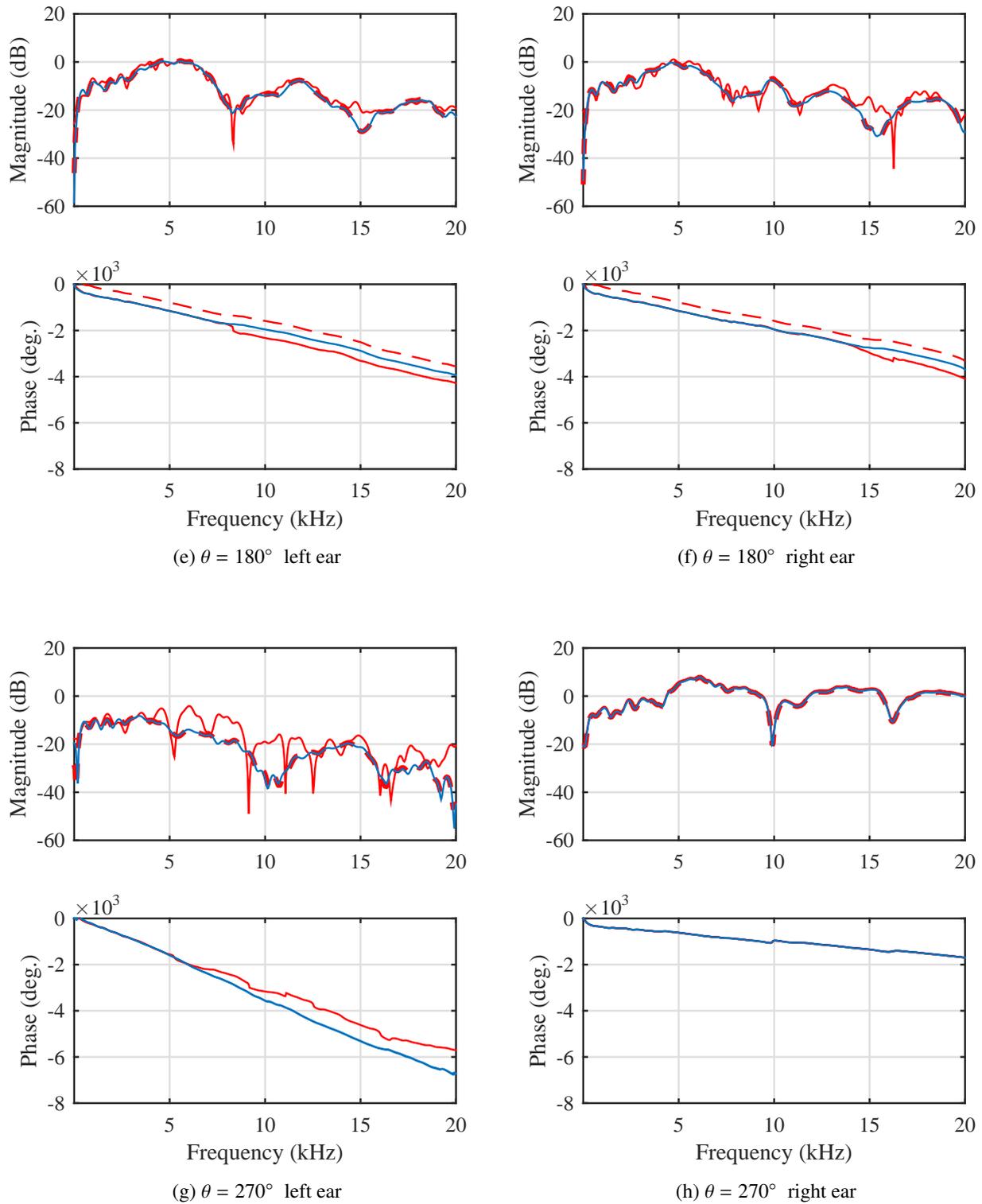


Fig. 3.10: Frequency characteristics of estimated HRIRs (red dashed line: direction-by-direction measurement, red solid line: simultaneous estimation with 520 sample shift, blue line: simultaneous estimation with 8,640 sample shift)

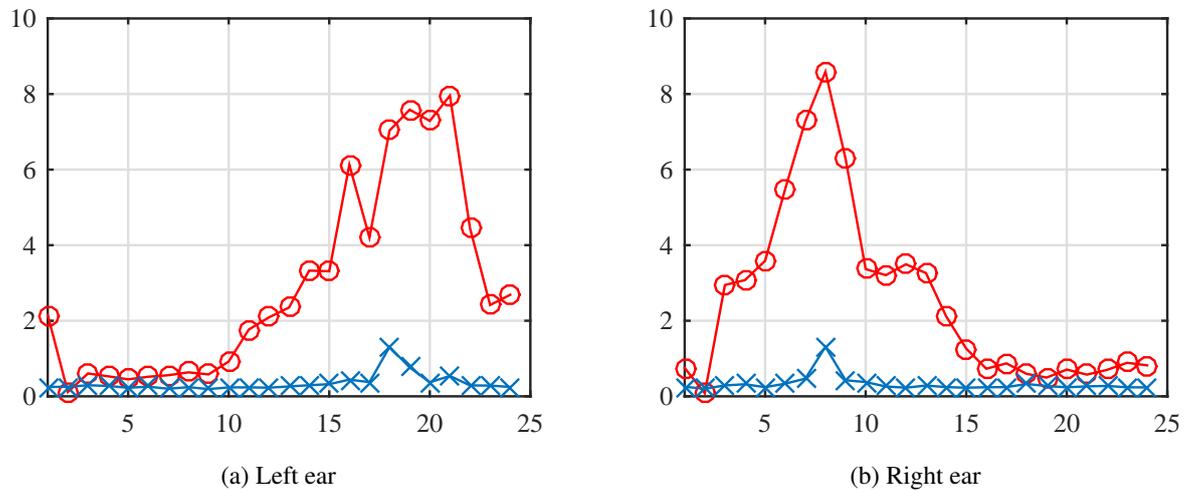


Fig. 3.11: SDs of estimated HRTFs (red solid line: simultaneous estimation with 520 sample shift, blue solid line: simultaneous estimation with 8,640 sample shift)

バルを 3 s としたため，測定時間は 171 s であった．同時推定に要した時間は約 0.68 s ( $\approx 2^{15} - 1/48,000$ ) であり，測定時間が大幅に短縮されたことが分かる．

提案法を用いて人の HRTF を実測する場合，同時再生される測定信号全体でその音圧レベルが調整されるため，十分な SN 比の確保が課題となっている．これに対して，Log-TSP を用いた 1 方向ずつ測定は，高い SN 比に加え，高調波歪みに代表される非線形歪みの分離が可能であることを特長としている．これは，測定信号を M 系列から作成する提案法と比較して，1 方向ずつの測定の優れている点である．このように，測定信号や測定法により特長は異なるが，測定を実施する実験室の暗騒音を下げ，電気音響変換器や電気音響機器の雑音対策を適切にとることができれば，測定時間を大幅に短縮できる提案法は非常に有効な手段である．



## 第 4 章

# 出力追従制御を応用したトランス オーラル再生制御器の設計

トランスオーラル再生法では、HRTF の個人性や不完全性による再生品質の低下、制御器の不確かさによるクロストーク・キャンセレーションの性能低下が問題となる。それらに加えて、制御器のゲインはその不確かさや入力信号に混入する雑音成分を増幅することから、この増幅によって受聴者に知覚される歪みやアーティファクトも問題となっている。

本章では、出力追従制御器を応用したトランスオーラル再生法を提案する。この制御器のゲイン抑制を目的として、状態空間でのトランスオーラル再生制御器の設計を、制御器の  $H_\infty$  ノルムを評価指標とする凸計画問題として定式化する。

### 4.1 出力追従制御を応用したトランスオーラル再生法

Fig. 4.1 にトランスオーラル再生の概略図を示す。図に示すように、2 次音源から制御点までのシステムを制御対象 (Plant)、仮想音源から制御点までのシステムを目標システム (Target) とする。第 2 章で述べたように、トランスオーラル再生は、この制御対象におけるクロストークを逆システムにより補償し、そこに所望信号を入力することにより実現される。Fig. 4.1 の場合、仮想音源からの再生を再現することから、目標システムの実出力が所望信号となる。

通常、制御対象はむだ時間を有する非最小位相系であり、不安定零点を含んでいる。そのため、制御器となる逆システムは不安定となる。従来法では、因果性を満たすために逆システムに遅延を付与する、正則化法や TSVD 法などの緩和処理を導入することによりその安定化を図ってきた。しかし、逆システムに付与される遅延はそのまま処理遅延として再生を遅らせ、緩和処理は安定性とトレードオフで逆システムの精度を低下させる。こ

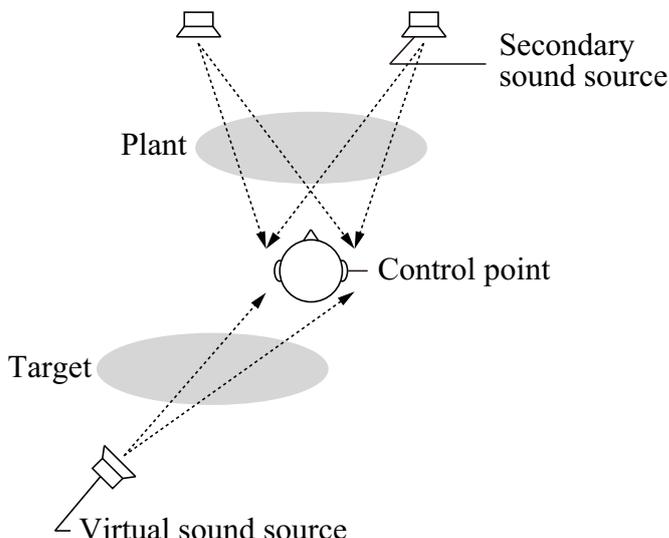


Fig. 4.1: Schematic diagram illustrating transaural reproduction

れらは，従来法における課題となっていた．

再度 Fig. 4.1 を見てみる．制御点において 2 次音源からの再生により仮想音源からの再生を再現するトランスオーラル再生は，制御対象の出力を目標システムの出力に追従させる出力追従制御とみなすことができる．この出力追従制御器，すなわち所望信号を出力する目標システムと制御対象の逆システムの直列結合系をトランスオーラル再生の制御器と考える．

制御対象の逆システムの構成には，非干渉化制御を援用する．非干渉化制御は，相互干渉の存在する MIMO システムにおいて，何らかの方法によりその相互干渉を取り除き，入出力を 1 対 1 に対応づける制御方式である．状態空間での非干渉化は出力フィードバックあるいは状態フィードバックによる伝達関数行列の対角化によって実現され [75–80]，移動体制御 [81]，プロセス制御 [82] など，様々な分野で応用されている [83–85]．

第 2 章で述べたように，トランスオーラル再生では，HRTF の個人性や不完全性，受聴環境の変化など，再生音場の不確かさによる再生品質の低下がしばしば問題となる．これに加えて，再生音場の不確かさは，その再生音場を制御対象として設計される制御器の不確かさとなる．この不確かさによりクロストークが効率的に抑制されなくなると，トランスオーラル再生自体が成立しなくなる．また，制御器のゲインはその不確かさや入力信号に混入する雑音成分を増幅する．増幅された不確かさや雑音成分は受聴者に歪みやアーティファクトとして知覚され，振幅の大きな周波数成分が含まれるような場合には，人の聴覚機構を損傷する危険もある．

提案法では，この制御器のゲイン抑制を目的として，状態空間でのトランスオーラル再生制御器の設計を，制御器の  $\mathcal{H}_\infty$  ノルムを評価指標とする凸計画問題として定式化する．

$\mathcal{H}_\infty$  ノルムは伝達関数行列の全周波数帯域における最大特異値であり，ゲインの上界を表し，再生音場の不確かさや入力信号に混入する雑音成分に対する抑制効果，すなわちロバスト性の指標となる関数である．提案法ではまず，非干渉化制御により制御対象の逆システムを構成する．ただし，音場再生ではフィードバック制御系の構築が困難なため，実装に際して等価なフィードフォワード制御系に変換する．次に，この逆システムと目標システムの直列結合系を構成し出力追従制御器とする．この出力追従制御器をトランスオーラル再生制御器とし，その安定性を制約として， $\mathcal{H}_\infty$  ノルムを最小化する状態フィードバックゲインを設計する．出力追従制御器を制御器とすることにより目標システムの特性能も含めて最適化が図られる．そのため，逆システムのみを制御器とする従来法と比較して，より高いロバスト性が得られると考える．また，制御対象と目標システムのむだ時間が相殺されるため，処理遅延の短い制御器として設計可能であることも特長として挙げられる．

## 4.2 線形行列不等式

提案法では，制約の記述に線形行列不等式 (Linear Matrix Inequality: LMI) を用いる．LMI とは，未知変数が行列に線形依存する不等式である [86, 87]．LMI における不等号は行列の定値性を表現し，未知変数  $x_i$  ( $i = 1, 2, \dots, m$ ) に関する LMI の一般形は

$$F(\mathbf{x}) := F_0 + \sum_{i=1}^m x_i F_i > 0 \quad (4.1)$$

により与えられる．ただし， $F_i$  ( $i = 0, 1, \dots, m$ ) は実対称行列であり， $\mathbf{x} = [x_1, x_2, \dots, x_m]^T$  は変数ベクトルである．

LMI 可解問題とは，式 (4.1) を満足する変数ベクトル  $\mathbf{x}$  を求める，もしくはそのような変数ベクトル  $\mathbf{x}$  が存在するか否かを判定する問題である．また，LMI 制約のもとで線形の目的関数を最小化 (最大化) する問題は半正定値計画問題 (Semidefinite Programming Problem: SDP) と呼ばれ，SDP の一般形は

$$\inf \mathbf{c}^T \mathbf{x} \text{ subject to } F(\mathbf{x}) \geq 0 \quad (4.2)$$

により与えられる．ただし， $\mathbf{c} \in \mathbb{R}^m$  は与えられた定数ベクトルである．内点法 [88] に代表される計算アルゴリズムの進展により，理論的に解析解を求めることが困難な問題に対し，十分な速度で数値解を与えることが可能となっている．

## 4.3 非干渉化制御による逆システムの構成

出力追従制御器の設計手続きを説明するに先立ち，本節では非干渉化制御による逆システムの構成について説明する [78, 80]．

対象として  $m$  入力  $p$  出力 ( $m > p$ ) の離散時間線形時不変システム  $G(z)$  を考える． $G(z)$  は 1 サンプル以上のむだ時間を含む，すなわち直達項を持たないものとし，その状態空間表現は

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \quad (4.3)$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) \quad (4.4)$$

と表される．ただし， $\mathbf{x}(k) \in \mathbb{R}^n$ ， $\mathbf{u}(k) \in \mathbb{R}^m$ ， $\mathbf{y}(k) \in \mathbb{R}^p$  はそれぞれ状態変数ベクトル，入力ベクトル，出力ベクトルであり， $\mathbf{A} \in \mathbb{R}^{n \times n}$ ， $\mathbf{B} \in \mathbb{R}^{n \times m}$ ， $\mathbf{C} \in \mathbb{R}^{p \times n}$  はシステムの係数行列を表す． $n$  はシステムの次数である．以降，このようなシステムを

$$\begin{aligned} G(z) &= \mathbf{C}(z\mathbf{I}_n - \mathbf{A})^{-1}\mathbf{B} \\ &= \left[ \begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{0}_{p \times m} \end{array} \right] \end{aligned} \quad (4.5)$$

と表記する．

また， $G(z)$  がむだ時間を含むことから，次式を満たすプロパーな遅れ型逆システム  $H_0(z)$  を考える．

$$\mathbf{G}(z)\mathbf{H}_0(z) = \mathbf{Q}(z) \quad (4.6)$$

ただし， $\mathbf{Q}(z)$  は対角成分に遅延要素を持つ伝達関数行列であり，遅延器として働く．

$$\mathbf{Q}(z) = \text{diag} \left[ z^{-\beta_1}, z^{-\beta_2}, \dots, z^{-\beta_p} \right] \quad (4.7)$$

$\beta_i$  ( $i = 1, 2, \dots, p$ ) はシステムの入力に信号を印加した時間を 0 としたとき  $i$  番目の出力において最初に非零の値が得られるサンプル時間であり，

$$\beta_i = \min(k | \mathbf{C}_i \mathbf{A}^{k-1} \mathbf{B} \neq \mathbf{0}_{1 \times m}; k = 1, \dots, N) \quad (4.8)$$

により与えられる．ただし，

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \vdots \\ \mathbf{C}_p \end{bmatrix}, \quad \mathbf{C}_i \in \mathbb{R}^{1 \times n} \quad (4.9)$$

であり， $N$  は入出力サンプル数である．

まず， $G(z)$  の非干渉化を考える． $G(z)$  に含まれるむだ時間は式 (4.7) のようになることから，出力を

$$\mathbf{y}(k) = \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_p(k) \end{bmatrix} \quad (4.10)$$

とおくと，時刻  $k + \beta_i$  での出力は

$$\begin{aligned} \begin{bmatrix} y_1(k + \beta_1) \\ y_2(k + \beta_2) \\ \vdots \\ y_p(k + \beta_p) \end{bmatrix} &= \begin{bmatrix} C_1 A^{\beta_1} \\ C_2 A^{\beta_2} \\ \vdots \\ C_p A^{\beta_p} \end{bmatrix} \mathbf{x}(k) + \begin{bmatrix} C_1 A^{\beta_1-1} \mathbf{B} \\ C_2 A^{\beta_2-1} \mathbf{B} \\ \vdots \\ C_p A^{\beta_p-1} \mathbf{B} \end{bmatrix} \mathbf{u}(k) \\ &= \mathbf{\Psi} \mathbf{x}(k) + \mathbf{\Phi} \mathbf{u}(k) \end{aligned} \quad (4.11)$$

となる．ただし，

$$\mathbf{\Psi} = \begin{bmatrix} C_1 A^{\beta_1} \\ C_2 A^{\beta_2} \\ \vdots \\ C_p A^{\beta_p} \end{bmatrix}, \quad \mathbf{\Phi} = \begin{bmatrix} C_1 A^{\beta_1-1} \mathbf{B} \\ C_2 A^{\beta_2-1} \mathbf{B} \\ \vdots \\ C_p A^{\beta_p-1} \mathbf{B} \end{bmatrix} \quad (4.12)$$

とおいた．このとき， $\mathbf{\Phi} \in \mathbb{R}^{p \times m}$  が行フルランクであれば，

$$\mathbf{\Phi} \hat{\mathbf{\Phi}} = I_p \quad (4.13)$$

を満たす  $\hat{\mathbf{\Phi}} \in \mathbb{R}^{m \times p}$  が存在する．この  $\hat{\mathbf{\Phi}}$  を使って  $G(z)$  に状態フィードバック

$$\mathbf{u}(k) = -\hat{\mathbf{\Phi}} \mathbf{\Psi} \mathbf{x}(k) + \hat{\mathbf{\Phi}} \mathbf{v}(k), \mathbf{v}(k) \in \mathbb{R}^p \quad (4.14)$$

を適用すると

$$\begin{aligned} \begin{bmatrix} y_1(k + \beta_1) \\ y_2(k + \beta_2) \\ \vdots \\ y_p(k + \beta_p) \end{bmatrix} &= \mathbf{\Psi} \mathbf{x}(k) + \mathbf{\Phi} \left\{ -\hat{\mathbf{\Phi}} \mathbf{\Psi} \mathbf{x}(k) + \hat{\mathbf{\Phi}} \mathbf{v}(k) \right\} \\ &= \mathbf{\Psi} \mathbf{x}(k) - \mathbf{\Psi} \mathbf{x}(k) + \mathbf{v}(k) \\ &= \mathbf{v}(k) \end{aligned} \quad (4.15)$$

となり，入出力が非干渉化される．また，式 (4.15) は

$$\mathbf{y}(k) = \mathbf{Q} \mathbf{v}(k) \quad (4.16)$$

と表すことができ，これを閉ループシステムとして表記すると

$$\left[ \begin{array}{c|c} \mathbf{A} - \mathbf{B} \hat{\mathbf{\Phi}} \mathbf{\Psi} & \mathbf{B} \hat{\mathbf{\Phi}} \\ \hline \mathbf{C} & \mathbf{0}_{p \times p} \end{array} \right] = \mathbf{Q}(z) \quad (4.17)$$

となる．式 (4.17) は状態フィードバックによる非干渉化を表している．

次に，式 (4.17) を等価なフィードフォワードシステムへ変換する．そのために

$$\mathbf{H}_0(z) = \left[ \begin{array}{c|c} \mathbf{A}_0 & \mathbf{B}_0 \\ \hline \mathbf{C}_0 & \mathbf{D}_0 \end{array} \right] \quad (4.18)$$

とし,  $G(z)$  と  $H_0(z)$  の直列結合系を考える.

$$\begin{aligned} G(z)H_0(z) &= \left[ \begin{array}{c|c} A & B \\ \hline C & \mathbf{0} \end{array} \right] \left[ \begin{array}{c|c} A_0 & B_0 \\ \hline C_0 & D_0 \end{array} \right] \\ &= \left[ \begin{array}{cc|c} A_0 & \mathbf{0}_{n \times n} & B_0 \\ \hline BC_0 & A & BD_0 \\ \hline \mathbf{0}_{p \times n} & C & \mathbf{0}_{p \times p} \end{array} \right] \end{aligned} \quad (4.19)$$

ただし,  $A_0 \in \mathbb{R}^{n \times n}$ ,  $B_0 \in \mathbb{R}^{n \times p}$ ,  $C_0 \in \mathbb{R}^{m \times n}$ ,  $D_0 \in \mathbb{R}^{m \times p}$  である. 変換行列

$$T = \left[ \begin{array}{cc} I_n & \mathbf{0}_{n \times n} \\ I_n & I_n \end{array} \right] \quad (4.20)$$

を使って等価変換を行い, 更に

$$A_0 = A + BC_0, \quad B_0 = BD_0 \quad (4.21)$$

とすると, 式 (4.19) は

$$G(z)H_0(z) = \left[ \begin{array}{c|c} A + BC_0 & BD_0 \\ \hline C & \mathbf{0}_{p \times p} \end{array} \right] \quad (4.22)$$

となる. 式 (4.6), (4.22) より

$$\left[ \begin{array}{c|c} A + BC_0 & BD_0 \\ \hline C & \mathbf{0} \end{array} \right] = Q(z) \quad (4.23)$$

となることから, 式 (4.23) は,  $G(z)$  に状態フィードバック

$$u(k) = C_0 x(k) + D_0 v(k) \quad (4.24)$$

を適用した閉ループシステムと解釈することができる. したがって, 閉ループシステムを表す式 (4.17), (4.23) を比較することにより

$$C_0 = -\hat{\Phi}\Psi, \quad D_0 = \hat{\Phi} \quad (4.25)$$

が導かれ, 更に

$$A_0 = A + BC_0, \quad B_0 = BD_0 \quad (4.21)$$

であることから, 式 (4.6) は

$$G(z) \left[ \begin{array}{c|c} A - B\hat{\Phi}\Psi & B\hat{\Phi} \\ \hline -\hat{\Phi}\Psi & \hat{\Phi} \end{array} \right] = Q(z) \quad (4.26)$$

と書き直される. すなわち,

$$\begin{aligned} H_0(z) &= \left[ \begin{array}{c|c} A_0 & B_0 \\ \hline C_0 & D_0 \end{array} \right] \\ &= \left[ \begin{array}{c|c} A - B\hat{\Phi}\Psi & B\hat{\Phi} \\ \hline -\hat{\Phi}\Psi & \hat{\Phi} \end{array} \right] \end{aligned} \quad (4.27)$$

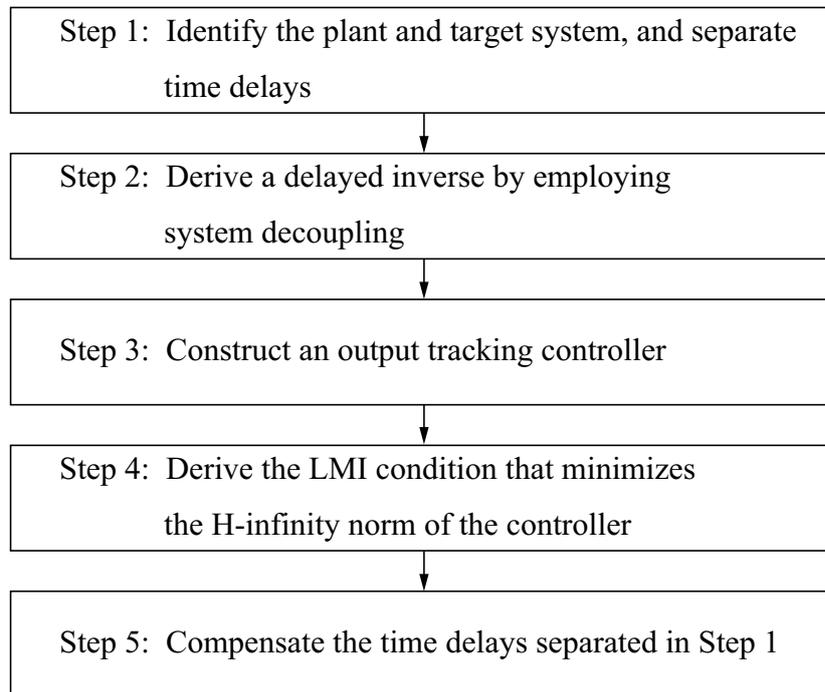


Fig. 4.2: Design procedure of a tracking controller

は， $G(z)$  のプロパーな遅れ型逆システムとなる．ただし，離散時間システムの因果性から，通常， $H_0(z)$  は安定なシステムとならないことに留意する．

#### 4.4 出力追従制御器の設計手続き

本節では，提案するトランスオーラル再生法に係る出力追従制御器の設計手続きを，Fig. 4.4 にしたがって説明する．まず Step 1 では，制御対象と目標システムを入出力データより同定し，遅延器による補償が容易なむだ時間を分離する．次に Step 2 では，4.3 節で説明した非干渉化制御により，制御対象の逆システムを構成する．ただし，この段階では状態フィードバックゲインを未定とする．続いて Step 3 では，この逆システムと目標システムとの直列結合系を構成する．この直列結合系が，出力追従制御のための制御器となる．したがって，その安定性と  $\mathcal{H}_\infty$  性能を特徴づける LMI 制約を導き，SDP として最適なフィードバックゲインを求めることが次の手続き Step 4 となる．最後に Step 5 では，Step 1 で分離したむだ時間を相殺して補償し，遅れを最小限とした出力追従制御器として構成する．

以降，制御対象を  $m_p$  入力  $p$  出力 ( $m_p > p$ ) の離散時間線形時不変システム  $G_p(z)$ ，目標システムを  $m_t$  入力  $p$  出力の離散時間線形時不変システム  $G_t(z)$  として，各手続きの詳細を説明する．

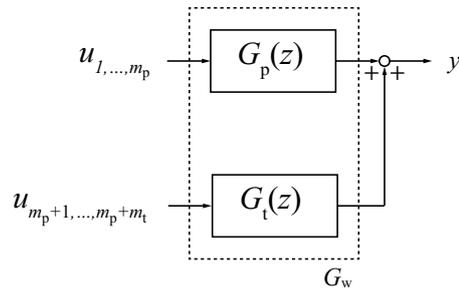


Fig. 4.3: Parallel structured system for system identification

### Step 1 システム同定とむだ時間の分離

Fig. 4.1 から分かるように，制御対象と目標システムは，出力端となる制御点を共有し，入力端をそれぞれの音源とする，再生音場全体のサブシステムと考えることができる．そこで，Fig. 4.3 に示すように， $G_p(z)$  と  $G_t(z)$  を出力端を共有する並列結合系として，入出力データより状態空間モデルとして同定する．同定法については特に定めないが，第3章で提案した同時推定法などによりインパルス応答を求め，特異値分解法 [71] を用いて実現すると都合がよい．なぜなら，このあと述べるむだ時間の分離をインパルス応答より直接行うことができるためである．特異値分解法の手続きについては，文献 [71] より付録 A にまとめる．

得られた状態空間モデルを  $m_p + m_t$  入力  $p$  出力の離散時間線形時不変システム  $G_w(z)$  とし，

$$G_w(z) = \left[ \begin{array}{c|c} A_w & B_w \\ \hline C_w & \mathbf{0}_{p \times (m_p + m_t)} \end{array} \right] \quad (4.28)$$

と表記する．ただし， $A_w \in \mathbb{R}^{n \times n}$ ， $B_w \in \mathbb{R}^{n \times (m_p + m_t)}$ ， $C_w \in \mathbb{R}^{p \times n}$  である．

まず，この  $G_w(z)$  より入力側のむだ時間を分離する．

$$G_w(z) = \tilde{G}_w(z) R_w(z) \quad (4.29)$$

ただし， $R_w(z)$  は，対角成分に遅延要素を持つ伝達関数行列である．

$$R_w(z) = \text{diag} \left[ z^{-\alpha_1 + 1}, z^{-\alpha_2 + 1}, \dots, z^{-\alpha_{m_p + m_t} + 1} \right] \quad (4.30)$$

$\alpha_i$  ( $i = 1, 2, \dots, m_p + m_t$ ) は  $G_w(z)$  の  $i$  番目の入力に信号を印加した時間を 0 としたとき出力において最初に非零の値が得られるサンプル時間であり，

$$\alpha_i = \min(k | C_w A_w^{k-1} B_{wi} \neq \mathbf{0}_{p \times 1}; k = 1, \dots, N) \quad (4.31)$$

により与えられる．ただし，

$$\mathbf{B}_w = \begin{bmatrix} \mathbf{B}_{w1} & \mathbf{B}_{w2} & \cdots & \mathbf{B}_{w(m_p+m_t)} \end{bmatrix}, \quad \mathbf{B}_{wi} \in \mathbb{R}^{n \times 1} \quad (4.32)$$

であり， $N$  は入出力サンプル数である．この  $\mathbf{R}_w(z)$  は，2 次音源および仮想音源のそれぞれから見た伝達関数に含まれるむだ時間である．

次に，

$$\tilde{\mathbf{G}}_w(z) = \left[ \begin{array}{c|c} \tilde{\mathbf{A}}_w & \tilde{\mathbf{B}}_w \\ \hline \tilde{\mathbf{C}}_w & \mathbf{0}_{p \times (m_p+m_t)} \end{array} \right]$$

とおき， $\tilde{\mathbf{G}}_w(z)$  より出力側のむだ時間を分離する．

$$\tilde{\mathbf{G}}_w(z) = \mathbf{Q}_w(z) \bar{\mathbf{G}}_w(z) \quad (4.33)$$

ただし，

$$\mathbf{Q}_w(z) = \text{diag} \left[ z^{-\beta_1+1}, z^{-\beta_2+1}, \dots, z^{-\beta_p+1} \right] \quad (4.34)$$

である． $\beta_i$  ( $i = 1, 2, \dots, p$ ) は， $\tilde{\mathbf{G}}_w(z)$  に信号を印加した時間を 0 としたとき  $i$  番目の出力において最初に非零の値が得られるサンプル時間であり，

$$\beta_i = \min(k | \tilde{\mathbf{C}}_{wi} \tilde{\mathbf{A}}_w^{k-1} \tilde{\mathbf{B}}_w \neq \mathbf{0}_{1 \times (m_p+m_t)}; k = 1, \dots, N) \quad (4.35)$$

により与えられる．ただし，

$$\tilde{\mathbf{C}}_w = \begin{bmatrix} \tilde{\mathbf{C}}_{w1} \\ \tilde{\mathbf{C}}_{w2} \\ \vdots \\ \tilde{\mathbf{C}}_{wp} \end{bmatrix}, \quad \tilde{\mathbf{C}}_{wi} \in \mathbb{R}^{1 \times n} \quad (4.36)$$

である．この  $\mathbf{Q}_w(z)$  は，式 (4.29) により音源から見たむだ時間を除いたのち制御点から見たときに，なお  $\mathbf{G}_w(z)$  に含まれるむだ時間となる．音源が受聴者の正中面より左右方向に分散配置されるような場合には，このむだ時間はなく  $\mathbf{Q}_w(z)$  は単位行列となる．音源が受聴者の正中面より片側に集中して配置されるような場合に， $\mathbf{Q}_w(z)$  は対角の伝達関数行列となる．

以上により，並列結合系  $\mathbf{G}_w(z)$  は

$$\mathbf{G}_w(z) = \mathbf{Q}_w(z) \bar{\mathbf{G}}_w(z) \mathbf{R}_w(z) \quad (4.37)$$

と表すことができる．このとき  $\mathbf{G}_p(z)$ ， $\mathbf{G}_t(z)$  は，そのサブシステムとして

$$\begin{aligned} \mathbf{G}_w(z) &= \begin{bmatrix} \mathbf{G}_p(z) & \mathbf{G}_t(z) \end{bmatrix} \\ &= \mathbf{Q}_w(z) \left[ \begin{array}{c|c} \bar{\mathbf{A}}_w & \bar{\mathbf{B}}_p \bar{\mathbf{B}}_t \\ \hline \bar{\mathbf{C}}_w & \mathbf{0} \end{array} \right] \left[ \begin{array}{c|c} \mathbf{R}_p(z) & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{R}_t(z) \end{array} \right] \end{aligned} \quad (4.38)$$

すなわち

$$\begin{aligned} \mathbf{G}_p(z) &= \mathbf{Q}_w(z) \bar{\mathbf{G}}_p(z) \mathbf{R}_p(z) \\ &= \mathbf{Q}_w(z) \left[ \begin{array}{c|c} \bar{\mathbf{A}}_w & \bar{\mathbf{B}}_p \\ \hline \bar{\mathbf{C}}_w & \mathbf{0} \end{array} \right] \mathbf{R}_p(z) \end{aligned} \quad (4.39)$$

$$\begin{aligned} \mathbf{G}_t(z) &= \mathbf{Q}_w(z) \bar{\mathbf{G}}_t(z) \mathbf{R}_t(z) \\ &= \mathbf{Q}_w(z) \left[ \begin{array}{c|c} \bar{\mathbf{A}}_w & \bar{\mathbf{B}}_t \\ \hline \bar{\mathbf{C}}_w & \mathbf{0} \end{array} \right] \mathbf{R}_t(z) \end{aligned} \quad (4.40)$$

により与えられる．ただし， $\bar{\mathbf{A}}_w \in \mathbb{R}^{n \times n}$ ， $\bar{\mathbf{B}}_p \in \mathbb{R}^{n \times m_p}$ ， $\bar{\mathbf{B}}_t \in \mathbb{R}^{n \times m_t}$ ， $\bar{\mathbf{C}}_w \in \mathbb{R}^{p \times n}$  であり，

$$\mathbf{R}_p(z) = \text{diag} \left[ z^{-\alpha_1+1}, z^{-\alpha_2+1}, \dots, z^{-\alpha_{m_p}+1} \right] \quad (4.41)$$

$$\mathbf{R}_t(z) = \text{diag} \left[ z^{-\alpha_{m_p+1}+1}, z^{-\alpha_{m_p+2}+1}, \dots, z^{-\alpha_{m_p+m_t}+1} \right] \quad (4.42)$$

である．

## Step 2 非干渉化制御による逆システムの構成

むだ時間を分離した制御対象  $\bar{\mathbf{G}}_p(z)$  より，4.3 節で示した非干渉化制御を援用してプロパーだが不安定な逆システム  $\mathbf{H}_0(z)$  を設計する．ただし，Step 1 でむだ時間を分離しているため，式 (4.7) における  $\beta_i$  ( $i = 1, 2, \dots, p$ ) はすべて 1 となる．したがって，式 (4.6) は

$$\bar{\mathbf{G}}_p(z) \mathbf{H}_0(z) = z^{-1} \mathbf{I}_p \quad (4.43)$$

となる．また，遅れ型逆システムを表す式 (4.27) は

$$\begin{aligned} \mathbf{H}_0(z) &= \left[ \begin{array}{c|c} \mathbf{A}_0 & \mathbf{B}_0 \\ \hline \mathbf{C}_0 & \mathbf{D}_0 \end{array} \right] \\ &= \left[ \begin{array}{c|c} \bar{\mathbf{A}}_w - \bar{\mathbf{B}}_p \hat{\Phi} \Psi & \bar{\mathbf{B}}_p \hat{\Phi} \\ \hline -\hat{\Phi} \Psi & \hat{\Phi} \end{array} \right] \end{aligned} \quad (4.44)$$

となる．ただし， $\beta_i = 1$  ( $i = 1, 2, \dots, p$ ) より

$$\Phi = \bar{\mathbf{C}}_w \bar{\mathbf{B}}_p, \quad \Psi = \bar{\mathbf{C}}_w \bar{\mathbf{A}}_w \quad (4.45)$$

である． $m_p > p$  であることから  $\hat{\Phi} \in \mathbb{R}^{m_p \times p}$  は縦長の矩形行列である．ここでは未知の変数行列とし，

$$\Phi \hat{\Phi} = \mathbf{I}_p \quad (4.46)$$

を満たすとき  $\bar{\mathbf{G}}_p(z)$  は非干渉化される．したがって，式 (4.46) は逆システムの成立する制約条件となる． $m_p > p$  であることから式 (4.46) は劣決定系となり， $\hat{\Phi}$  は解に自由度を持つ．

## Step 3 出力追従制御器の構成

むだ時間を分離した目標システム  $\bar{G}_t(z)$  と遅れ型逆システム  $H_0(z)$  との直列結合系を構成し,

$$\begin{aligned}
 H(z) &= H_0(z)\bar{G}_t(z) \\
 &= \left[ \begin{array}{c|c} \bar{A}_w - \bar{B}_p \hat{\Phi} \Psi & \bar{B}_p \hat{\Phi} \\ \hline -\hat{\Phi} \Psi & \hat{\Phi} \end{array} \right] \left[ \begin{array}{c|c} \bar{A}_w & \bar{B}_t \\ \hline \bar{C}_w & \mathbf{0} \end{array} \right] \\
 &= \left[ \begin{array}{cc|c} \bar{A}_w - \bar{B}_p \hat{\Phi} \Psi & \bar{B}_p \hat{\Phi} \bar{C}_w & \mathbf{0} \\ \mathbf{0} & \bar{A}_w & \bar{B}_t \\ \hline -\hat{\Phi} \Psi & \hat{\Phi} \bar{C}_w & \mathbf{0} \end{array} \right] \quad (4.47)
 \end{aligned}$$

とおく．この  $H(z)$  が，目標システムの出力に制御対象の出力を追従させるための出力追従制御器となる．したがって， $H(z)$  の安定化と  $\mathcal{H}_\infty$  ノルムを最小化する変数行列  $\hat{\Phi}$  を探索することが次の手続きとなる．

LMI に基づく制御器設計では，制御仕様に対応する条件式に代数的な操作を施すことにより，求解可能な LMI に帰着させる必要がある．式 (4.47) も，このまま LMI 条件を適用すると，双線形項を含み数値的求解が困難となる．また，変数行列  $\hat{\Phi}$  が式 (4.46) を満たすとき  $H_0(z)$  は  $\bar{G}_p(z)$  の遅れ型逆システムとなるが， $\bar{G}_p(z)$  は非最小位相系であり不安定零点を含むため，式 (4.46) の制約条件そのままでは  $H_0(z)$  の安定解を得ることができない．そのため，LMI 導出に係る前処理として，以下の式変換を行う．

まず， $\bar{G}_t(z)$  の係数行列  $\bar{A}_w$  が正則であることを仮定し，変換行列

$$T = \begin{bmatrix} \bar{A}_w^{-1} & \mathbf{0}_{n \times n} \\ \mathbf{I}_n & \mathbf{I}_n \end{bmatrix} \quad (4.48)$$

を使って等価変換を行う．この変換により，直列結合される  $\bar{G}_t(z)$  と  $H_0(z)$  の間で，すなわち制御器の内部で極と零点が相殺され，式 (4.47) は

$$\begin{aligned}
 H(z) &= \left[ \begin{array}{cc|c} \bar{A}_w & \bar{A}_w \bar{B}_p \hat{\Phi} \bar{C}_w & \mathbf{0} \\ \mathbf{0} & -\bar{A}_w \bar{B}_p \hat{\Phi} \bar{C}_w + \bar{A}_w & \bar{B}_t \\ \hline \mathbf{0} & \hat{\Phi} \bar{C}_w & \mathbf{0} \end{array} \right] \\
 &= \left[ \begin{array}{c|c} -\bar{A}_w \bar{B}_p \hat{\Phi} \bar{C}_w + \bar{A}_w & \bar{B}_t \\ \hline \hat{\Phi} \bar{C}_w & \mathbf{0} \end{array} \right] \quad (4.49)
 \end{aligned}$$

となる．

次に、式 (4.49) では  $\hat{\Phi}$  は常に  $\bar{C}_w$  との積として現れることに注目する．これは  $\bar{G}_t(z)$  と  $H_0(z)$  の直列結合系として  $H(z)$  を構成したことにより現れた項である．このとき

$$\Phi \hat{\Phi} = I_p \quad (4.46)$$

は常に  $\bar{C}_w$  との積として評価されることから

$$\Phi \hat{\Phi} \bar{C}_w = \bar{C}_w \quad (4.50)$$

として問題はなく、更に、新たな変数行列を  $\bar{\Phi} := \hat{\Phi} \bar{C}_w \in \mathbb{R}^{m_p \times n}$  と定義することにより、

$$\Phi \bar{\Phi} = \bar{C}_w \quad (4.51)$$

と変換することができる．式 (4.46) が満たされるとき式 (4.50) も常に満たされることから、式 (4.46) は式 (4.51) の成立するための十分条件であるといえる．一方で、式 (4.51) が成立するためには、必ずしも式 (4.46) が満たされる必要はない．このことから、式 (4.51) は式 (4.46) から緩和された出力追従制御の成立する制約条件と考えることができる．

同様の変数変換により、出力追従制御器  $H(z)$  は

$$H(z) = \left[ \begin{array}{c|c} -\bar{A}_w \bar{B}_p \bar{\Phi} + \bar{A}_w & \bar{B}_t \\ \hline \bar{\Phi} & \mathbf{0} \end{array} \right] \quad (4.52)$$

となる．

#### Step 4 LMI に基づく出力追従制御器の設計

式 (4.52) により表される出力追従制御器  $H(z)$  の安定化と  $\mathcal{H}_\infty$  ノルムを最小化する変数行列  $\bar{\Phi}$  を決定する．離散時間線形時不変システムの安定性と  $\mathcal{H}_\infty$  性能を特徴づける LMI 表現は、以下の定理により与えられる [87]．

## 定理 4.1

次式で表される直達項のない離散時間線形時不変システムを考える．

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \quad (4.53)$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) \quad (4.54)$$

このシステムの伝達関数を  $G(z)$  とする．このとき，与えられた  $\gamma > 0$  に対して，次の条件は等価である．

1. 行列  $\mathbf{A}$  がシュール安定であり， $\|G(z)\|_\infty < \gamma$  が成り立つ．
2. 次の LMI を満足する  $\mathbf{X} \in \mathbb{S}_{++}^n$  が存在する．

$$\begin{bmatrix} -\mathbf{X} + \mathbf{A}\mathbf{X}\mathbf{A}^\top + \mathbf{B}\mathbf{B}^\top & \mathbf{A}\mathbf{X}\mathbf{C}^\top \\ \mathbf{C}\mathbf{X}\mathbf{A}^\top & \mathbf{C}\mathbf{X}\mathbf{C}^\top - \gamma^2 \mathbf{I} \end{bmatrix} < 0 \quad (4.55)$$

なお， $\|\cdot\|_\infty$  は  $\mathcal{H}_\infty$  ノルムを表す．

定理 4.1 より， $G(z)$  の  $\mathcal{H}_\infty$  ノルムは

$$\inf_{\mathbf{X} \in \mathbb{S}_{++}^n} \gamma_{\text{sq}} \quad \text{subject to} \quad \begin{bmatrix} -\mathbf{X} + \mathbf{A}\mathbf{X}\mathbf{A}^\top + \mathbf{B}\mathbf{B}^\top & \mathbf{A}\mathbf{X}\mathbf{C}^\top \\ \mathbf{C}\mathbf{X}\mathbf{A}^\top & \mathbf{C}\mathbf{X}\mathbf{C}^\top - \gamma_{\text{sq}} \mathbf{I} \end{bmatrix} < 0 \quad (4.56)$$

を解くことで計算できる．この SDP の最適値を  $\gamma_{\text{sq}}^*$  とすれば， $\|G(z)\|_\infty = \sqrt{\gamma_{\text{sq}}^*}$  となる．

式 (4.55) の LMI は，シュールの補題 [86] より

$$\begin{bmatrix} -\mathbf{X} & * & * & * \\ \mathbf{0} & -\gamma^2 \mathbf{I} & * & * \\ \mathbf{X}\mathbf{A}^\top & \mathbf{X}\mathbf{C}^\top & -\mathbf{X} & * \\ \mathbf{B}^\top & \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix} < 0 \quad (4.57)$$

と等価である．ただし，記述を簡単にするため，ブロック対称となる部分行列を ‘\*’ で表している．

この式 (4.57) に式 (4.52) を代入することにより，

$$\begin{bmatrix} -\mathbf{X} & * & * & * \\ \mathbf{0} & -\gamma^2 \mathbf{I} & * & * \\ \mathbf{X}(-\bar{\mathbf{A}}_w \bar{\mathbf{B}}_p \bar{\Phi} + \bar{\mathbf{A}}_w)^\top \mathbf{X} \bar{\Phi}^\top - \mathbf{X} & * & * & * \\ \bar{\mathbf{B}}_t^\top & \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix} < 0 \quad (4.58)$$

が導かれる．式 (4.58) には  $\mathbf{X}$  と変数行列  $\bar{\Phi}$  に関する双線形項が含まれるが，新たな変数行列を  $\mathbf{Y} := \bar{\Phi} \mathbf{X} \in \mathbb{R}^{m_p \times n}$  と定義することにより，等価的に

$$\begin{bmatrix} -\mathbf{X} & * & * & * \\ \mathbf{0} & -\gamma^2 \mathbf{I} & * & * \\ (-\bar{\mathbf{A}}_w \bar{\mathbf{B}}_p \mathbf{Y} + \bar{\mathbf{A}}_w \mathbf{X})^\top & \mathbf{Y}^\top & -\mathbf{X} & * \\ \bar{\mathbf{B}}_t^\top & \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix} < 0 \quad (4.59)$$

が導かれる．同様に，出力追従制御の制約条件となる式 (4.51) は， $Y$  を用いて

$$\Phi Y = \tilde{C}_w X \quad (4.60)$$

のように修正される．安定かつ  $\mathcal{H}_\infty$  ノルムを最小化する出力追従制御器  $H(z)$  の設計は，この式 (4.59)，(4.60) を連立させることにより SDP として定式化することができる．式 (4.59)，(4.60) の LMI が成り立つとき

$$\bar{\Phi} = YX^{-1} \quad (4.61)$$

となることから，この  $\bar{\Phi}$  を式 (4.52) に代入することにより  $H(z)$  を決定することができる．

### Step 5 むだ時間の補償

式 (4.38) で分離したむだ時間を，遅延器により補償する．ただし，出力側のむだ時間  $Q_w(z)$  は制御対象と目標システムとで共通となり相殺されるため，その補償処理は省略することができる．

制御対象における入力側のむだ時間は，2 次音源から制御点へ音響信号が伝搬する際の音波面の到達時間の遅れである．逆システムによりこの遅れを補償しようとするとき，音波面を進めるような非因果的な操作が必要となる．一方，目標システムにおいても，仮想音源から制御点への音波の伝搬に際して同じように到達時間の遅れが生じる．そのため，双方のむだ時間を相殺し，その差分を遅延器により補償する．

補償のための遅延器を

$$R_c(z) = \text{diag} [z^{-\gamma_1}, z^{-\gamma_2}, \dots, z^{-\gamma_{m_p}}] \quad (4.62)$$

とし，制御対象のむだ時間  $R_p(z)$  との直列結合系が目標システムのむだ時間  $R_t(z)$  のうち最小のものと等しくなるよう，その遅延量  $\gamma_i$  ( $i = 1, 2, \dots, m_p$ ) を決定する．更に，非干渉化制御の際に生じる 1 サンプル分の遅延 (式 (4.43)) も併せて  $R_c(z)$  で補償することにする．したがって，その関係式は

$$R_p(z)R_c(z) \left\{ z^{\alpha_i - 1} I_{m_p} \right\} = z I_{m_p} \quad (4.63)$$

のように与えられる．ただし，

$$\alpha_t = \min(\alpha_k; k = m_p + 1, m_p + 2, \dots, m_p + m_t) \quad (4.64)$$

である．

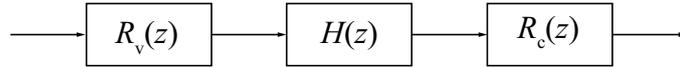


Fig. 4.4: Output tracking controller as a transaural reproduction controller

$R_p(z)$  のむだ時間が  $\alpha_t$  より大きいとき，すなわち 2 次音源からの音波面が仮想音源からの音波面より遅れて到達するような音源配置のときには，式 (4.63) を満たす  $R_c(z)$  はインプロパーとなる．この場合には，式 (4.63) を

$$R_p(z)R_c(z) \{z^{\alpha_t-1} I_{m_p}\} = z^{1-\alpha_c} I_{m_p} \quad (4.65)$$

のように修正する．式 (4.65) を満たし，かつ  $R_c(z)$  がプロパーとなるような最小の  $\alpha_c$  を求める．この  $\alpha_c$  が出力追従制御に係る遅れ時間となる．

併せて，仮想音源間のむだ時間を補償する遅延器を

$$R_v(z) = \text{diag} [z^{-\gamma_{m_p+1}}, z^{-\gamma_{m_p+2}}, \dots, z^{-\gamma_{m_p+m_t}}] \quad (4.66)$$

とし，その遅延量  $\gamma_i$  ( $i = m_p + 1, m_p + 2, \dots, m_p + m_t$ ) を  $R_t(z)$  と  $\alpha_c$  との差分として決定する．したがって，その関係式は

$$R_v(z) = R_t(z) \{z^{\alpha_t} I_{m_t}\} \quad (4.67)$$

のように与えられる．この遅延器は出力追従制御器  $H(z)$  の前段に配置され，仮想音源間の距離差により生じる音波面の到達時間の差を再現するものである．

以上の手続きにより導出された  $R_v(z)$  と  $H(z)$ ， $R_c(z)$  の直列結合系が，遅れを最小限する出力追従制御器，すなわちトランスオーラル再生制御器の構成となる (Fig. 4.4)．なお，提案法の有効性は，第 5 章のシミュレーション実験で検証する．



## 第 5 章

# 22.2 マルチチャンネル音響のトランスオーラル再生

22.2 マルチチャンネル音響（以下，22.2ch 音響）方式は，空間的に配置された 22 チャンネルと低音効果（Low Frequency Effects: LFE）用の 2 チャンネルから構成される，チャンネルベースの 3 次元音響方式である．2018 年に開始が予定されている 8K SHV の音響方式であり，その実用化に向けた研究，開発と国際規格化が進められている．

本章では，22.2ch 音響方式を概説したのち，提案するトランスオーラル再生法の 22.2ch 音響再生への適用について述べる．8K SHV は多様な視聴スタイルが予想されるが，多くの場合，家庭で 24 台のスピーカを配置することは困難であると考えられる．そこで，本章の前半では，ディスプレイに一体化される枠型スピーカアレイとラインアレイスピーカを用いた，22.2ch 音響のトランスオーラル再生法を提案する．本章の後半では，同スピーカを用いたシミュレーション実験により，第 3 章と第 4 章で提案した HRTF の多方向同時推定法とトランスオーラル再生法の有効性を検証する．

## 5.1 22.2 マルチチャンネル音響

### 5.1.1 22.2 マルチチャンネル音響方式の概要

8K SHV は，3,300 万画素の超高精細映像と 3 次元音響から成る次世代高臨場感放送システムである [89,90]．そのフォーマットは ITU-R，米国映画テレビ技術者会（Society of Motion Picture and Television Engineers: SMPTE），電波産業会（Association of Radio Industries and Businesses: ARIB）などの標準化機関において規格化が進められおり，2016 年に衛星による試験放送が開始された．今後，2018 年には実用放送の開始が，東京五輪の開催年となる 2020 年には，全国各地でのパブリックビューイングなど，本格普及に向

けた施策が予定されている。

8K SHV の音響方式として開発が進められる 22.2ch 音響方式は、視聴者を取り囲むように配置されたフルレンジの 22 チャンネルと、LFE 用の 2 チャンネルから構成され、その場にいるかのような高い臨場感と 3 次元的な音響空間構築を実現することができる方式である [91]。8K SHV の映像に対応する音響方式の要求条件は、ITU-R BS.1909 に以下のように規定される [92,93]。

#### 音質要求

1. The sound image should be reproduced in all directions around the listener, including the elevation direction, within reasonable limits of stability.  
(受聴者の周囲、仰角方向を含む全方向から、安定して音像が再生されること)
2. The sensation of a three-dimensional spatial impression that augments a sense of reality, which is related to ambience and envelopment, should be significantly enhanced over established sound formats in Recommendation ITU-R BS.775.  
(アンビエンスと包み込まれ感に関連し、現実感を増大させる 3 次元的な音響空間印象が、既存の音声フォーマットから著しく向上されること)
3. For applications with accompanying picture, the directional stability of the frontal sound image should be maintained over the entire area of high-resolution large-screen digital imagery. The coincidence of position between sound images and video images should also be maintained over a wide image and listening area.  
(映像を伴うアプリケーションにおいて、高精細大画面映像上の全領域における前方音像の方向が安定であること。広い視聴範囲において、音像と映像が一致すること)
4. Excellent sound quality should be maintained over a wide viewing/listening area.  
(広い視聴範囲において、高音質であること)
5. For the coding parameters of sound signals, see Recommendation ITU-R BS.646.  
(音響信号のコーディングパラメータに関しては、ITU-R BS.646 を参照のこと)
6. For the relative timing of sound and vision, see Recommendation ITU-R BT.1359. Implementers should note that in large venues where sound and vision timing varies relative to seating location, other adjustments may need to be taken into account.  
(映像と音響の相対的なタイミングに関しては、ITU-R BT.1359 を参照のこと。映像と音響のタイミングが座席により変化する大きな会場では、他の調整も考慮する必要があることに留意すること)

### 機能要求

1. Backward compatibility with the 5.1 channel sound system and conventional two-channel sound system should be ensured to an acceptable degree.  
(既存の 5.1 チャンネル音響方式および 2 チャンネルステレオ音響方式との後方互換性を有すること)
2. Appropriate rendering of the sound format should be ensured for loudspeaker configurations suitable for the range of typical listening environments, to provide an optimum sound field and maintain excellent sound quality at least within a limited area, or alternatively for headphone reproduction.  
(最適な音場を提供し、少なくとも限定的な範囲やヘッドホン再生に対してその高音質を維持するため、様々な視聴環境とスピーカ配置に対して音声フォーマットの適切なレンダリングを用意すること)

### 運用要求

1. Live recording, mixing and transmission should be possible, if required.  
(ライブ収録、ミクシングおよび伝送が可能であること)

これらの要求条件を満たし、既存のチャンネルベース音響方式のチャンネル配置を包含する 22.2ch 音響方式は、8K SHV の音響方式として十分な性能を有しているといえる。3 層からなるチャンネル配置を Fig. 5.1 に、チャンネル番号とラベル、名称の一覧を Table 5.1 に示す。ディスプレイに正対したとき視聴者の耳の高さとなる中層には 10 チャンネル、頭上を含む上層には 9 チャンネル、ディスプレイの下方となる下層には前方のみの 3 チャンネルが配置される。LFE チャンネルは特に配置を定めていないが、下層の左側の範囲に LFE-1 が、右側の範囲に LFE-2 が配置されることが多い。これらのチャンネル配置や名称は、国際的には Rec. ITU-R BS.2051 [94] に、国内では ARIB STD-B59 [95] に規定されている。

#### 5.1.2 22.2 マルチチャンネル音響の家庭での再生

8K SHV の家庭での視聴スタイルは、視聴する人数や部屋の広さ、ディスプレイ、音響再生システム等の構成要素によって異なってくる [96]。

Fig. 5.2 はそれぞれ、(a) シアタールーム、(b) リビングルーム、そして (c) プライベートルームでの視聴のイメージを示している。シアタールームは、高品質の映像、音響を視聴するために設計される。部屋の施工段階から防音工事やスピーカの埋め込み、そのための配線工事が可能であり、22.2ch 音響の再生システム設置に制約のない、自由度の高い部屋

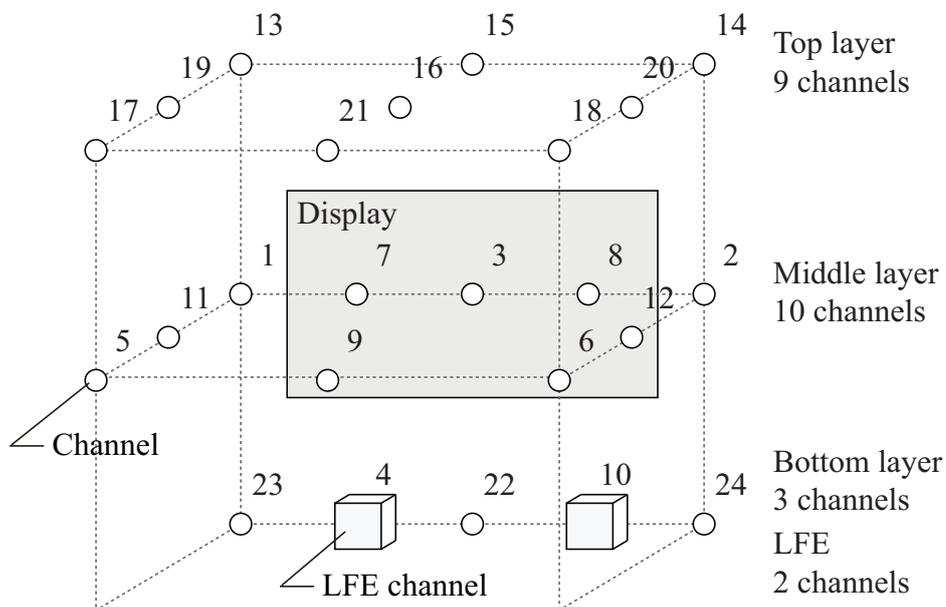


Fig. 5.1: Audio channel arrangement of 22.2 multichannel sound system

である。リビングルームは、1人での視聴や家族での視聴など、多様な状況を想定した部屋である。ディスプレイの大画面化によりリビングシアターも身近になってきたが、部屋の形状や大きさ、家具の配置などの事情から、22.2ch音響の再生システム設置にはある程度の制約が伴ってくる。プライベートルームでの視聴は人数も限られてくることから、ディスプレイもそれに応じて小型のものが選ばれる。22.2ch音響も、取り扱いの容易さから、据え置きやディスプレイ内蔵の少数スピーカによる再生が主流になると考えられる。

Fig. 5.3は、携帯端末での視聴のイメージを示している。携帯端末での視聴では、22.2ch音響の再生システムにも受聴環境を選ばない自由度が求められる。特に屋外での視聴においては、遮音性が高く、携帯性にも優れたイヤフォン、ヘッドフォンがこれまでと同様その中心になると考えられる。一方、屋内では、タブレット端末などでの視聴機会も増えると予想される。

## 5.2 ディスプレー一体型スピーカ

5.1節で示したように、8K SHVは多様な視聴スタイルが予想されることから、それぞれのスタイルに応じた22.2ch音響の再生法を提案する必要がある。多くの場合、家庭で24台のスピーカを配置することは困難である。したがって、22.2ch音響を、より少ないスピーカ数で、より簡易に受聴することができる再生法を8K SHV視聴の選択肢として提案することは、重要な課題の一つである。そこで、フラットパネルディスプレイ(Flat

Table 5.1: Channel number, label, and name of 22.2 multichannel sound system

No.	Label	Channel name	No.	Label	Channel name
1	FL	Front Left	13	TpFL	Top Front Left
2	FR	Front Right	14	TpFR	Top Front Right
3	FC	Front Centre	15	TpFC	Top Front Centre
4	LFE-1	Low Frequency Effects-1	16	TpC	Top Centre
5	BL	Back Left	17	TpBL	Top Back Left
6	BR	Back Right	18	TpBR	Top Back Right
7	FLc	Front Left centre	19	TpSiL	Top Side Left
8	FRc	Front Right centre	20	TpSiR	Top Side Right
9	BC	Back Centre	21	TpBC	Top Back Centre
10	LFE-2	Low Frequency Effects-2	22	BtFC	Bottom Front Centre
11	SiL	Side Left	23	BtFL	Bottom Front Left
12	SiR	Side Right	24	BtFR	Bottom Front Right

Panel Display: FPD) のベゼル<sup>\*1</sup>に外接して一体化されるスピーカと，このスピーカを用いたトランスオーラル再生法を提案する．

### 5.2.1 枠型スピーカアレイ

85 型 FPD のために試作した枠型 72 スピーカアレイを Fig. 5.4 に示す．枠型 72 スピーカアレイの 4 辺には計 72 個のスピーカユニットが等間隔で配置されている．これは，様々なスピーカユニット配置を想定して実験を実施するためであり，次節以降で示すシミュレーション実験での測定にはこのスピーカアレイを使用する．枠型 12 スピーカアレイを Fig. 5.5 に示す．枠型 12 スピーカアレイは，22.2ch 音響のトランスオーラル再生のために試作した，ディスプレイ一体型のスピーカアレイである．その上辺と下辺には各 5 個のスピーカユニットが等間隔で配置され，左右の側辺にはその中間位置に各 1 個のスピーカユニットが配置されている．この配置は，22.2ch 音響方式の前方チャンネルに対応する．枠上の各スピーカユニットは，相互の干渉と混変調を防ぐために，独立したキャビネットに収められている．それぞれのスピーカユニットは高い耐入力と低い歪み特性を有しており，振動板の直径は 7 cm と小サイズであるにもかかわらず，最大音圧レベルは

\*1 ディスプレイの外枠部分



(a) Theater room



(b) Living room



(c) Private room

Fig. 5.2: Concept images of viewing the 8K SHV in home

92 dB である。更に，Fig. 5.6 に示す振動板のエッジ構造は NHK の標準モニタ用の技術を小型化して導入したものであり，大振動時の高い直進性と逆共振の抑制を実現することにより，従来の同口径スピーカユニットと比べ，中域で 20 dB 程度の歪み低減に成功している。枠型 12 スピーカアレイの左右側辺にはそれぞれ 2 個のサブウーファユニットが備えられており，LFE チャンネルの再生により広がり感や包まれ感などの空間印象を向上させると共に，22.2ch 音響の低域成分の再生により，低域音響特性を改善している。

85 型 FPD に一体化される枠型デジタルスピーカアレイを Fig. 5.7 に示す。前出の枠型 12 スピーカアレイと同様，独立した 12 のキャビネットのそれぞれにスピーカユニットが収められている。デジタルアンプは高い電力効率と低い発熱量を特長としており，アナログアンプと比較して小型化，軽量化が容易である。試作したスピーカもスピーカユニットと同じキャビネットにデジタルアンプを内蔵することにより，コンパクトな外観と軽量化を実現している。振動板の直径は 7.7 cm，最大音圧レベルは 86 dB である。最低共振周波数が 100 Hz とやや高いため，ベースマネージメントにより低周波数帯域を補っている。



Fig. 5.3: Concept images of viewing the 8K SHV with mobile terminals

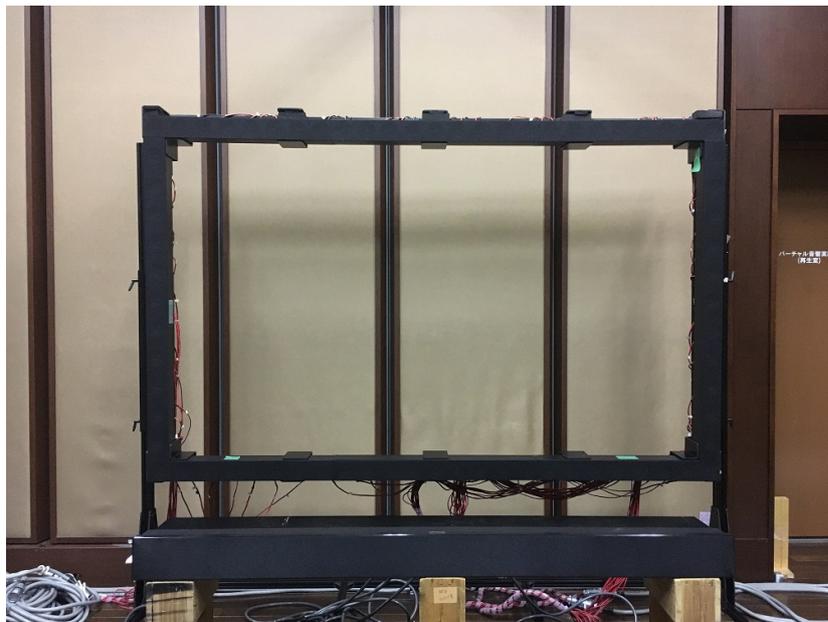


Fig. 5.4: 72-loudspeaker frame jointly developed by NHK and Foster Electric Company, Ltd.

### 5.2.2 ラインアレイスピーカ

ディスプレイの4辺にスピーカユニットを配置した枠型スピーカは、そのデザイン性と設置の難しさが、実用化に向けた課題の一つとなっている。

枠型デジタルスピーカアレイと同様デジタルアンプを内蔵したデジタルラインアレイスピーカを Fig. 5.8 に示す。ベゼルの上下に外接するラインアレイスピーカには各9個のスピーカユニットが等間隔で配置されている。スピーカユニットの特性は枠型デジタルス



Fig. 5.5: 12-loudspeaker frame integrated into an 85-inch LCD jointly developed by NHK and Foster Electric Company, Ltd.



Fig. 5.6: Corrugated design of diaphragm

ピーカアレイと同様である。ラインアレイスピーカは内蔵，外付けともに市場に流通しており，消費者にも受け入れられやすいと考える。

### 5.2.3 22.2 マルチチャンネル音響トランスオーラル再生

枠型 12 スピーカアレイによる再生では，22.2ch 音響方式の各チャンネルのうち前方チャンネルは，枠型 12 スピーカ上の各スピーカユニットに割り当てられる。ただし，ディスプレイの位置に重なる FLc チャンネル，FRc チャンネル，FC チャンネルは，振幅パン



Fig. 5.7: Digital loudspeaker frame jointly developed by NHK and Sharp Corp.

ニングにより仮想音像として合成される。前方音像の定位感を評価した聴取実験によれば、水平方向の振幅パニングと比較して、垂直方向の振幅パニングは前方により安定した仮想音像を合成することが可能である [97]。したがって、上記 3 チャンネルの合成には、各チャンネルを挟む上辺と下辺のスピーカユニットを用いている。側方、後方のチャンネルは、12 個すべてのスピーカユニットを 2 次音源としたトランスオーラル再生により、規格上の所定方向に音像として合成される。

ラインアレイスピーカによる再生では、FLc チャンネル、FRc チャンネル、FC チャンネルに加え、FL チャンネル、FR チャンネルも垂直方向の振幅パニングにより仮想音像として合成される。また、側方、後方のチャンネルのうち、中層に位置する 5 チャンネルは下辺のラインアレイスピーカを 2 次音源としたトランスオーラル再生により、上層に位置する 6 チャンネルは上辺のラインアレイスピーカを 2 次音源としたトランスオーラル再生により、規格上の所定方向に音像として合成される。第 2 章で示したように、主なスペクトルキューは音の入射角に依存して構成される。このことを考慮して、仮想音像に近い仰角を持つスピーカユニットを 2 次音源としている。



Fig. 5.8: Digital line array loudspeaker jointly developed by NHK and Sharp Corp.

### 5.3 実験 1: 枠型スピーカアレイによるシミュレーション 実験

第 4 章で提案したトランスオーラル再生法をディスプレイ体型スピーカによる 22.2ch 音響再生に適用し，シミュレーション実験によりその有効性を検証した．本節ではまず，枠型スピーカアレイを用いたシミュレーション実験の結果を示す．再生音場のインパルス応答は，第 3 章で提案した多方向同時推定法と，Log-TSP を測定信号として用いた 1 方向ずつの測定により取得した．

#### 5.3.1 実験条件

##### I. 再生音場の推定

再生音場の測定は，Fig. 5.9 に示す音源と受聴点配置のもと，NHK 放送技術研究所の音響無響室で実施した．この受音点配置は，22.2ch 音響の側方，後方チャンネルに相当する．測定の様子を Fig. 5.10 に示す．前述のように，測定には枠型 72 スピーカアレイを使用した．図中，HATS の正面に配置された枠型スピーカアレイがトランスオーラル再生における 2 次音源となり，HATS の立ち位置が受聴点，すなわちその両耳位置が制御点となる．側方，後方に配置されたスピーカは，仮想音源から制御点までの伝達関数を測定する

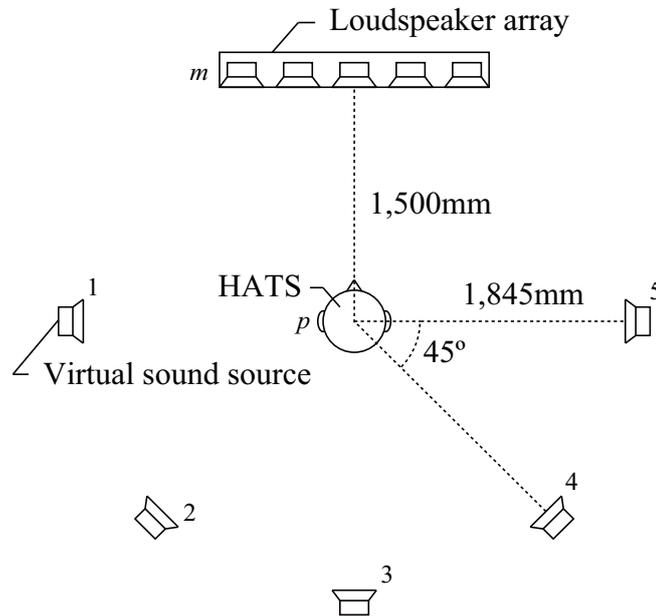


Fig. 5.9: Layout of sound sources and listening position

Table 5.2: Measurement signals in experiment 1

	MLS	Log-TSP
Sampling frequency (kHz)	48	48
Bit depth (bits)	24	24
Signal length (sample)	$2^{17} - 1$	$2^{17}$
Cyclic shift length (sample)	1,200	-
Model order	1,024	-
Sound pressure level (dB, A-weighted)	70	70

ための音源である．第 4 章で述べたように，この仮想音源から制御点までのシステムが目標システムとなり，2 次音源から制御点までのシステムが制御対象となる．

測定信号を Table 5.2 にまとめる．同時推定では，シフトレジスタ数 17 の M 系列より測定信号を作成した．このとき信号長は  $2^{17} - 1$  サンプルとなる．サンプリング周波数は 48 kHz，量子化ビット数は 24 ビットとした．FIR モデルの次数は 1,024 次とし，巡回シフト幅は 1,200 サンプルとした．測定信号を 2 次音源と仮想音源位置のスピーカより同時に再生し，その応答を HATS の左右耳位置で収録した．提案法により FIR モデルのパラメータを推定したのち，スピーカの逆特性，および 20 Hz から 20 kHz を通過帯域とするバンドパスフィルタを畳み込んだ信号をインパルス応答とした．

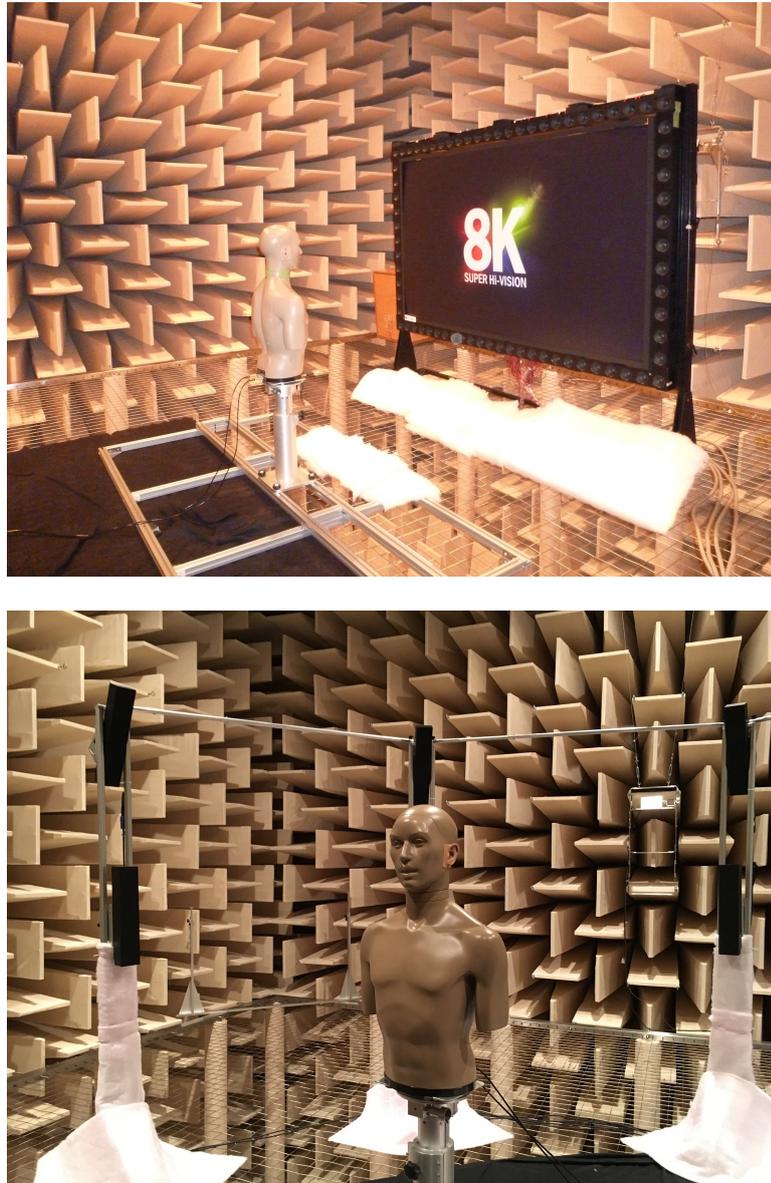


Fig. 5.10: Measurement of the reproduced sound field in an acoustic anechoic chamber of the NHK Science and Technology Research Laboratories

1 方向ずつの測定では，信号長  $2^{17}$  サンプルの Log-TSP を測定信号とした．サンプリング周波数，量子化ビット数は M 系列と同じ 48 kHz，24 ビットとした．2 次音源と仮想音源位置のスピーカのそれぞれより測定信号を再生し，HATS の左右耳位置で収録された応答に逆 Log-TSP，スピーカの逆特性，および 20 Hz から 20 kHz を通過帯域とするバンドパスフィルタを畳み込んだ信号をインパルス応答とした．

測定音圧は HATS の頭部位置においてピークを 70 dB (A-weighted) とし，1 方向ずつの測定では，応答長を 1,024 サンプルとして矩形窓で切り出した．ただし，応答には暗騒

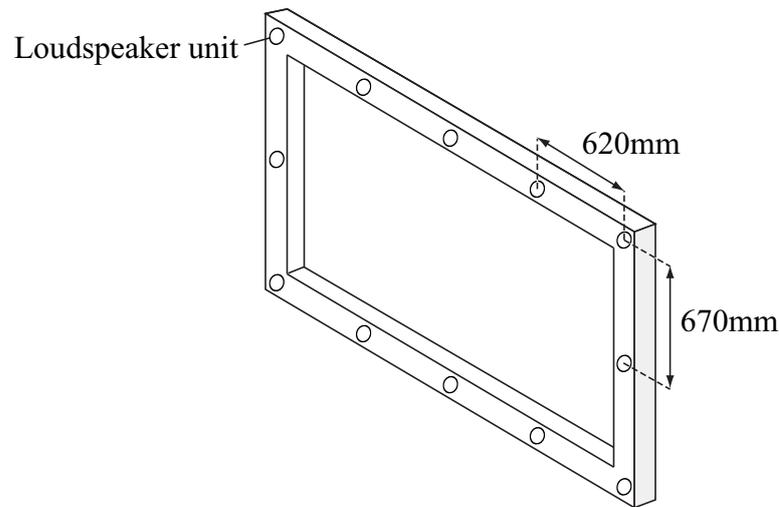


Fig. 5.11: Layout of the loudspeaker units in experiment 1

音などの外乱入力が含まれるため，全応答のピーク振幅から  $-32$  dB を閾値とし，各インパルス応答の先頭から振幅が閾値以上となるまでの区間，およびピークから減衰し振幅が閾値以下となって以降の区間は無音区間とした．このように測定されるインパルス応答を HRIR，その伝達関数表現を HRTF と呼ぶことも多いが，スピーカの指向特性や伝搬路での遅延も応答に含まれることから，ここでは単に再生音場のインパルス応答，伝達関数と呼ぶことにする．

## II. 出力追従制御器の設計

測定したインパルス応答から，Fig. 5.9 に示す各仮想音源に対して，その出力を追従する制御器を設計する．Fig. 5.11 に実験 1 で 2 次音源として使用したスピーカユニットの配置を示す．受聴点は Fig. 5.9 に示す 1 点とした．したがって，制御対象は 12 入力 2 出力の，目標システムは 1 入力 2 出力の離散時間線形時不変システムとなる．

インパルス応答から状態空間モデルの実現には特異値分解法を用いた．このとき，提案法の Step 1 にしたがって分離されるむだ時間は，インパルス応答先頭の無音区間から直接推定することが可能である．また，モデルの係数行列  $A$  は正則となる．モデルの次数は，HRIR の一般的な次数にならい 256 次とした．

LMI/SDP の MATLAB 上でのプログラミングには YALMIP (Yet Another LMI Parser) [98] を，SDP ソルバには SeDuMi [99] を用いた．

Table 5.3: SDs generated by modeling error (simultaneous estimation, dB)

Control point	1	2	3	4	5
Left ear	1.35	1.55	1.81	3.07	3.41
Right ear	7.88	2.61	2.44	1.95	1.82

### 5.3.2 実験結果

Fig. 5.9 上で 4 とラベル付けされた受聴点より右後方の仮想音源を例に，その結果を周波数特性図で示す．なお，ここでは便宜上，むだ時間を分離したのち状態空間モデルに変換した目標システムを目標モデル，制御対象を制御対象モデルと呼ぶことにする．また，同時推定により得られたインパルス応答を用いた実験を同時推定に基づく実験，1 方向ずつの測定により得られたインパルス応答を用いた実験を 1 方向ずつの測定に基づく実験と呼ぶことにする．

#### I. 再生音場のモデル化精度

まず，むだ時間を分離した目標システムのインパルス応答と，そのインパルス応答から実現した目標モデルの，それぞれの周波数特性を Figs. 5.12–5.13 に示す．Fig. 5.12 は同時推定により得られたインパルス応答からの実現を示し，Fig. 5.13 は 1 方向ずつの測定により得られたインパルス応答からの実現を示す．図中，赤実線はリファレンスとなる目標システムのインパルス応答の周波数特性であり，青実線は目標モデルの周波数特性である．いずれも図も，応答全体の振幅包絡および位相特性は，おおよそ近似できていることを示している．しかし，モデルの次数を 256 次に制限しているため，特に 14 kHz を超える高周波数帯域で，モデル化誤差により生じる歪みが増加している．

この歪みを SD を用いて定量化し，制御点ごと，仮想音源ごとの一覧としてまとめたのが Tables 5.3–5.4 である．Table 5.3 は同時推定に基づく実験の結果を示し，Table 5.4 は 1 方向ずつの測定に基づく実験の結果を示す．いずれの表も，仮想音源に対して同側耳側の制御点では歪みが小さく抑えられているが，頭部の陰となる対側耳側の制御点では拡大する傾向を示している．なお，同時推定に基づく実験での平均 SD は 2.79 dB，1 方向ずつの測定に基づく実験での平均 SD は 2.18 dB であった．

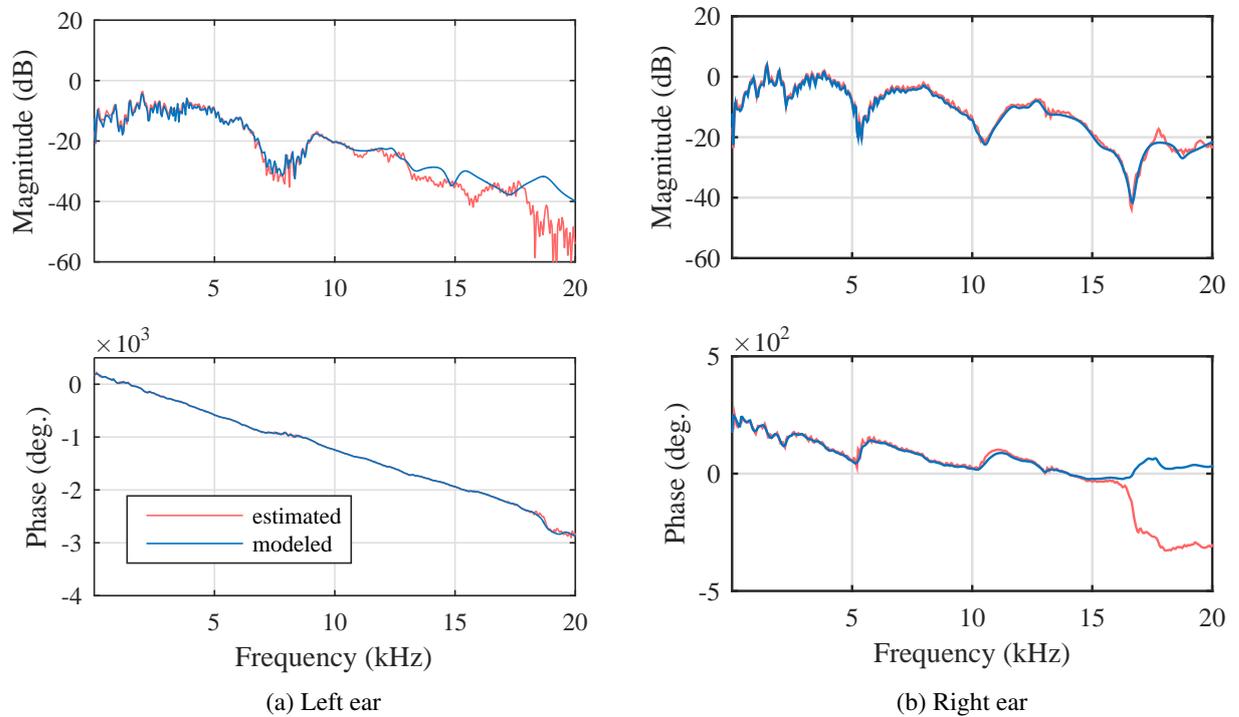


Fig. 5.12: Comparison of frequency characteristics (simultaneous estimation, red solid line: measured impulse responses, blue solid line: target model)

Table 5.4: SDs generated by modeling error (direction-by-direction measurement, dB)

Control point	1	2	3	4	5
Left ear	1.28	1.11	1.57	2.35	3.29
Right ear	3.92	1.47	2.44	2.86	1.47

## II. 出力追従制御器の設計精度

次に，設計した出力追従制御器と制御対象モデルの直列結合系の周波数特性を Figs. 5.14–5.15 に示す．これは，仮にむだ時間がないものとしたとき，制御点で合成される応答である．リファレンスは目標モデルの周波数特性とした．Fig. 5.14 は同時推定に基づく実験の結果を示し，Fig. 5.15 は 1 方向ずつの測定に基づく実験の結果を示す．図中，赤実線はリファレンスとなる目標モデルの周波数特性であり，青実線は直列結合系の周波数特性である．Fig. 5.15 は，わずかな位相遅れを除いて，目標モデルの特性が忠実

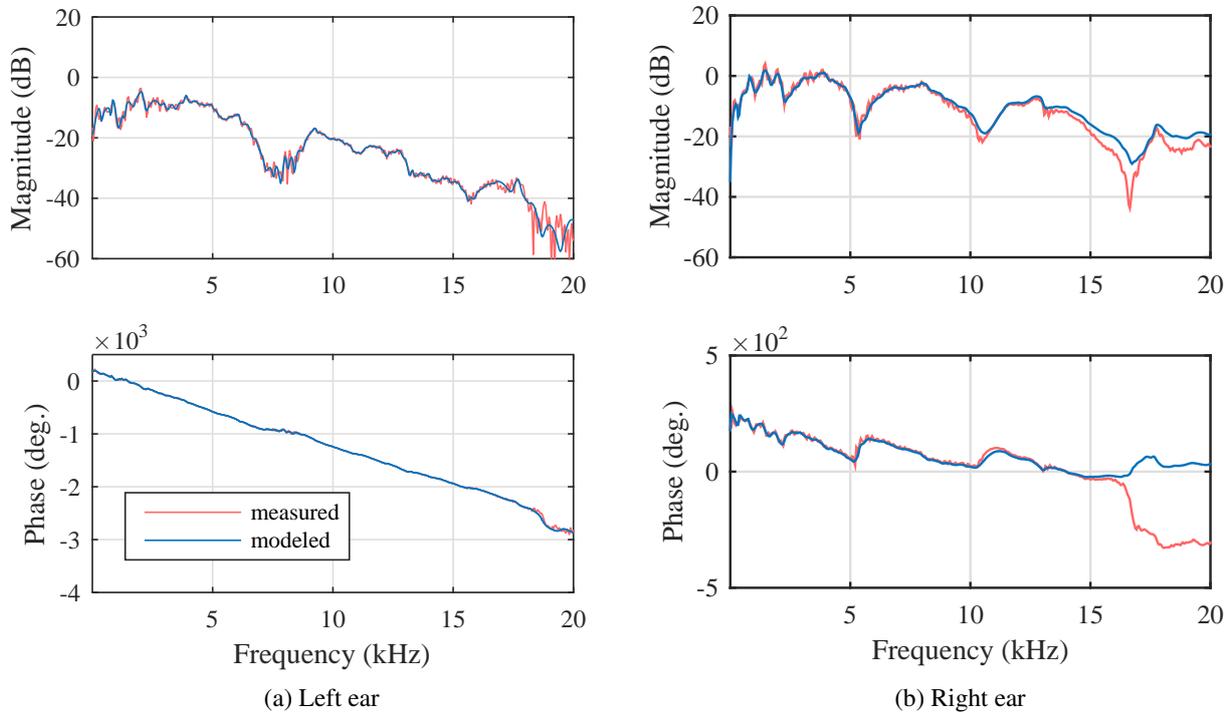


Fig. 5.13: Comparison of frequency characteristics (direction-by-direction measurement, red solid line: measured impulse responses, blue solid line: target model)

に再現されていることを示している．この位相遅れは，提案法の Step 2，式 (4.43) に示す 1 サンプル分の遅延に相当する．すなわち，理論的には提案法によりほぼ理想的な出力追従制御が可能であることを示している．一方 Fig. 5.14 では，Fig. 5.15 と比較して，位相遅れが大きくなっているように見える．これはプロット時の位相アンラップによる誤差の累積が原因であり，図より，直流に近い低周波数帯域で発生していることが分かる．その点を除いて，これもほぼ理想的な出力追従制御が可能であることを示している．

Tables 5.5–5.6 に，仮想音源ごとに設計される出力追従制御器の  $\mathcal{H}_\infty$  ノルムを示す．Table 5.5 は同時推定に基づく実験の結果を示し，Table 5.6 は 1 方向ずつの測定に基づく実験の結果を示す．参考として，提案法の Step 2，式 (4.46) の  $\hat{\Phi}$  として Moore-Penrose の擬似逆行列を用いた場合も併記した．いずれの表も，提案法により，すべての仮想音源に対して制御器の  $\mathcal{H}_\infty$  ノルムが大幅に小さくなることを示している．無音区間判定の閾値，モデルの次数などによりその大きさは変動するが，擬似逆行列を用いた場合と比較してより小さな値を返す傾向は変わらなかった．また，インパルス応答の推定法，仮想音源によって， $\mathcal{H}_\infty$  ノルムの大きさにばらつきがあることが分かる．これは，測定法の差よりむしろ，無音区間を判定する閾値処理，特にインパルス応答の立ち上がりを判定する閾値処

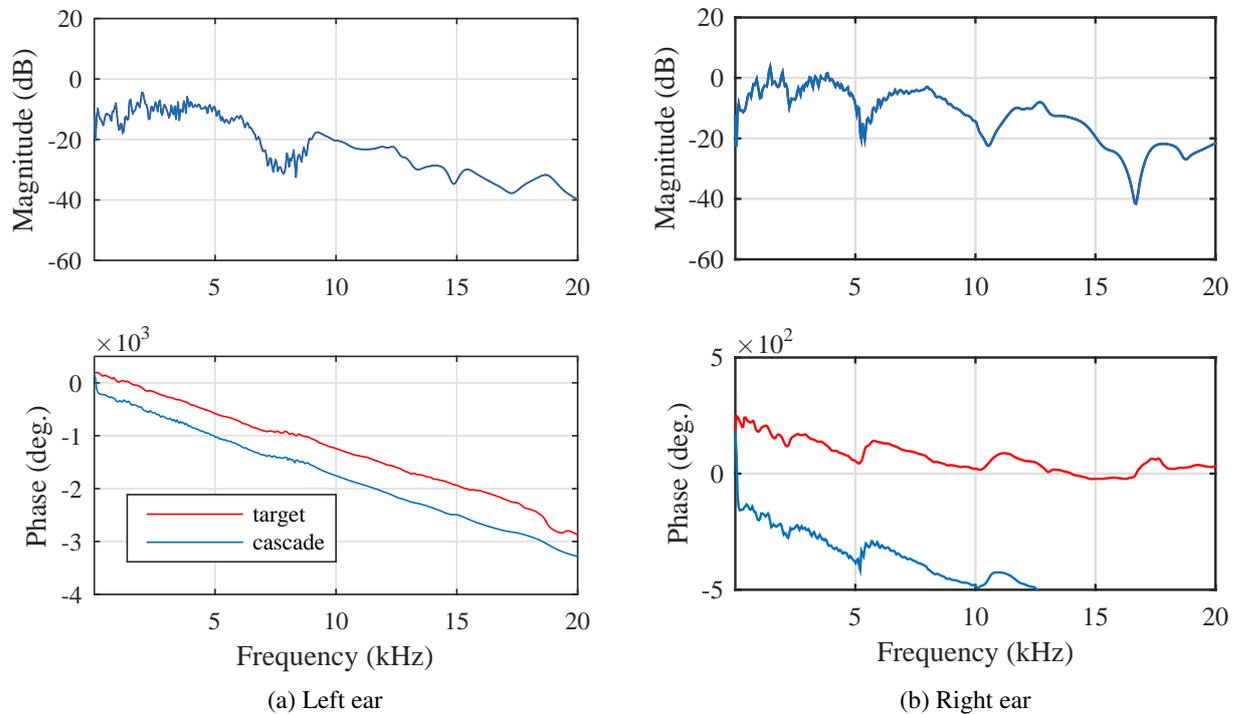


Fig. 5.14: Comparison of frequency characteristics (simultaneous estimation, red solid line: target model, blue solid line: cascade structured system)

Table 5.5:  $\mathcal{H}_\infty$  norm of the tracking controllers (simultaneous estimation)

Method	1	2	3	4	5
Proposed	3.01	5.81	2.55	6.30	4.37
Moore-Penrose	7.88	60.85	6.32	24.50	30.07

理に起因する。

### III. 出力追従制御器の再生精度

最後に、実際のトランスオーラル再生を想定し、設計した出力追従制御器のインパルス応答に制御対象のインパルス応答を畳み込んだ応答の周波数特性を Figs. 5.16–5.17 に示す。Fig. 5.16 は同時推定に基づく実験の結果を示し、Fig. 5.17 は 1 方向ずつの測定に基づく実験の結果を示す。図中、赤実線はリファレンスとした目標システムのインパルス応答から求めた周波数特性であり、青実線は畳み込み出力、すなわち出力追従制御出力の周

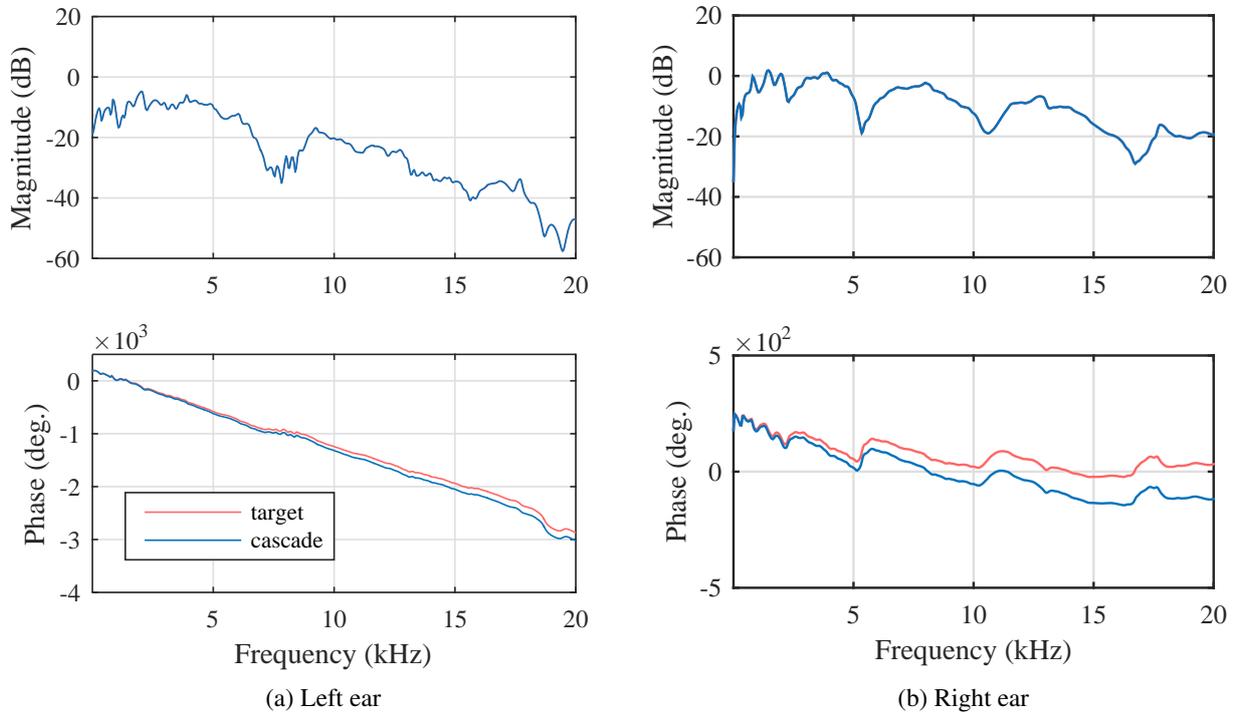


Fig. 5.15: Comparison of frequency characteristics (direction-by-direction measurement, red solid line: target model, blue solid line: cascade structured system)

Table 5.6:  $\mathcal{H}_\infty$  norm of the tracking controllers (direction-by-direction measurement)

Method	1	2	3	4	5
Proposed	3.49	7.91	3.34	1.45	3.62
Moore-Penrose	12.35	76.99	6.87	16.67	29.16

波数特性である．いずれの図も，応答全体の振幅包絡や位相特性，振幅特性上のスペクトルキューなど，音像の定位知覚に係る手掛かりはおおよそ再現されていることを示している．しかし，目標システムのモデル化誤差を評価した Figs. 5.12–5.13 と比較して，明らかに歪みが増加する結果となった．これは，目標システムに加え，制御対象のモデル化誤差も累積されるためである．

制御点ごと，仮想音源ごとの SD を Tables 5.7–5.8 にまとめる．Table 5.7 は同時推定に基づく実験の結果を示し，Table 5.8 は 1 方向ずつの測定に基づく実験の結果を示す．いずれの表からも，仮想音源の位置が正中面から離れるに従い，歪みが増加していることが

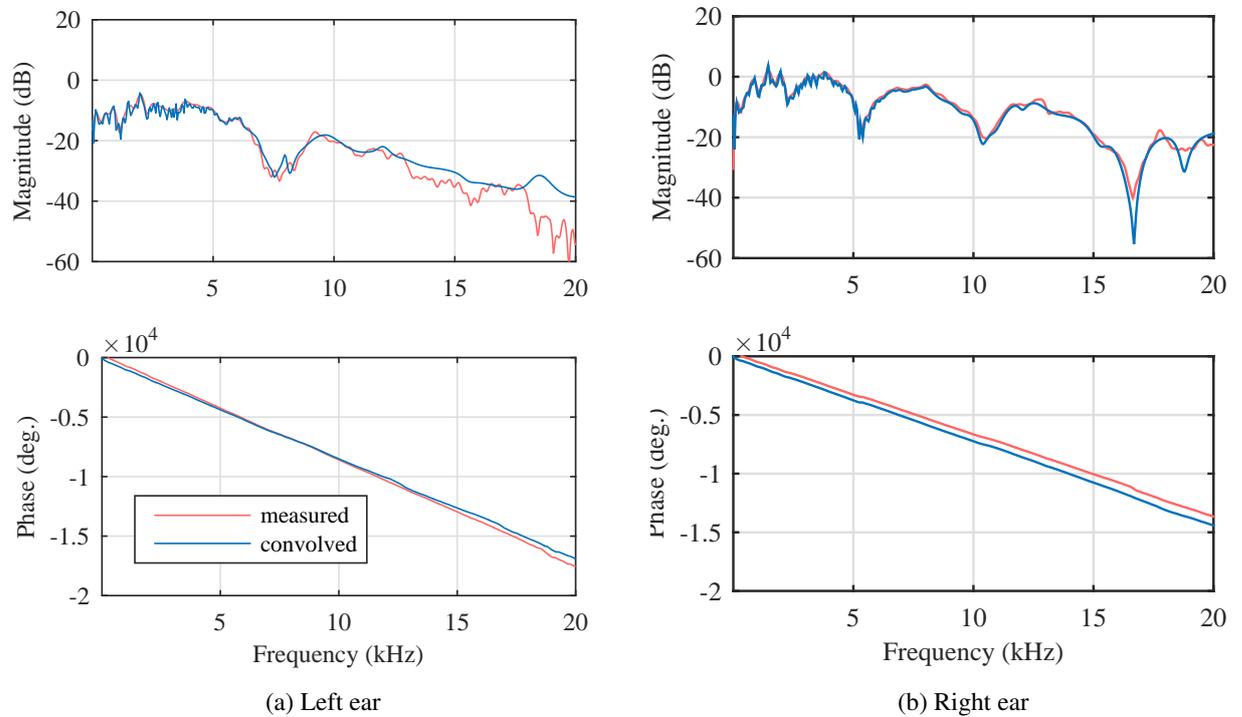


Fig. 5.16: Comparison of frequency characteristics (simultaneous estimation, red solid line: measured, blue solid line: convolved)

Table 5.7: SDs at control points (simultaneous estimation, dB)

Control point	1	2	3	4	5
Left ear	6.10	2.87	2.97	4.94	7.12
Right ear	10.09	4.12	2.88	4.75	5.94

分かる．両耳間時間差が大きくなると，むだ時間を分離したのちにも振幅の小さな応答区  
間が残ることになる．この区間は信号対雑音比が低く，モデル化誤差増加の主要因となっ  
ていると考える．

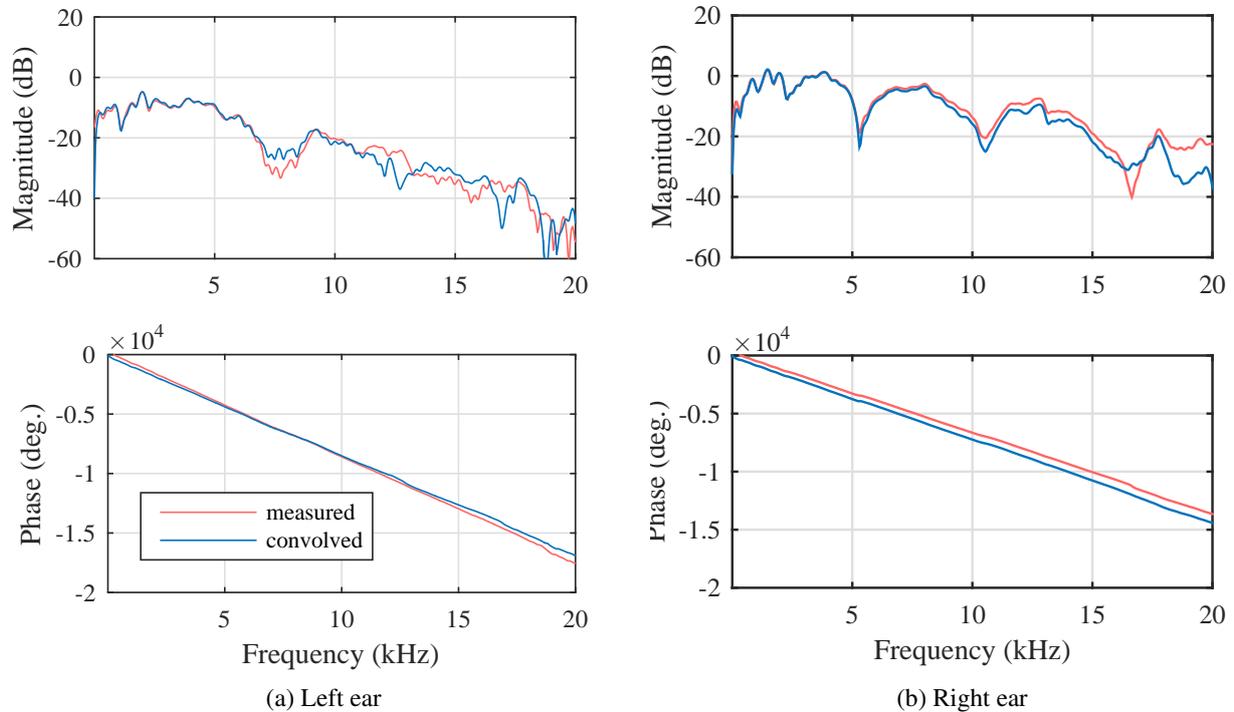


Fig. 5.17: Comparison of frequency characteristics (direction-by-direction measurement, red solid line: measured, blue solid line: convolved)

Table 5.8: SDs at control points (direction-by-direction measurement, dB)

Control point	1	2	3	4	5
Left ear	6.20	2.28	2.29	4.11	6.63
Right ear	8.67	3.45	2.94	3.64	4.55

## 5.4 実験 2: ラインアレイスピーカによるシミュレーション 実験

本節では、ラインアレイスピーカを用いたシミュレーション実験の結果を示す。再生音場のインパルス応答は、Log-TSP を測定信号として用いた 1 方向ずつの測定により取得した。

Table 5.9: Measurement signals in experiment 2

	Log-TSP
Sampling frequency (kHz)	48
Bit depth (bits)	24
Signal length (sample)	$2^{17}$
Sound pressure level (dB, A-weighted)	70

### 5.4.1 実験条件

#### I. 再生音場の測定

再生音場の測定は、実験 1 と同じ Fig. 5.9 に示す音源と受音点配置のもと、NHK 放送技術研究所の音響無響室で実施した。ただし、測定には Fig. 5.10 に示す枠型 72 スピーカアレイの下辺のみを取り外して使用した。

測定信号を Table 5.9 にまとめる。実験 2 では 1 方向ずつの測定のみとし、信号長  $2^{17}$  サンプルの Log-TSP を測定信号とした。サンプリング周波数、量子化ビット数は実験 1 と同じ 48 kHz、24 ビットとした。2 次音源と仮想音源位置のスピーカのそれぞれより測定信号を再生し、HATS の左右耳位置で収音した応答に逆 Log-TSP、スピーカの逆特性、および 20 Hz から 20 kHz を通過帯域とするバンドパスフィルタを畳み込んだ信号をインパルス応答とした。

測定音圧は HATS の頭部位置においてピークを 70 dB (A-weighted) とし、応答長を 1,024 サンプルとして矩形窓で切り出した。実験 1 と同じく、全応答のピーク振幅から  $-32$  dB を閾値とし、各インパルス応答の先頭から振幅が閾値以上となるまでの区間、およびピークから減衰し振幅が閾値以下となって以降の区間は無音区間とした。

#### II. 出力追従制御器の設計

測定したインパルス応答から、Fig. 5.9 に示す各仮想音源に対して、その出力を追従する制御器を設計する。Fig. 5.18 に実験 2 で 2 次音源として使用したスピーカユニットの配置を示す。受聴点は実験 1 と同じ 1 点とした。したがって、制御対象は 13 入力 2 出力の、目標システムは 1 入力 2 出力の離散時間線形時不変システムとなる。

インパルス応答から状態空間モデルへの変換には、実験 1 と同じ特異値分解法を用いた。分離するむだ時間は、インパルス応答先頭の無音区間から直接推定した。このとき、閾値処理により判定する無音区間が実験 1 と比較して長くなったため、モデルの次数は

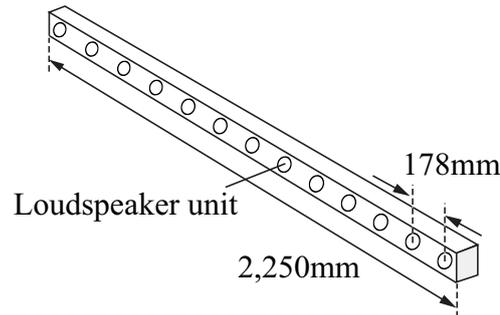


Fig. 5.18: Layout of the loudspeaker units in experiment 2

128 次とした．棒型 72 スピーカレイの下辺を除いて撤去した結果，治具での反射が実験 1 より減少したためである．

#### 5.4.2 実験結果

Fig. 5.9 上で 4 とラベル付された受聴点より右後方の仮想音源を例に，その結果を周波数特性図で示す．

##### I. 再生音場のモデル化精度

まず，むだ時間を分離した目標システムの実測インパルス応答と，そのインパルス応答から実現した目標モデルの，それぞれの周波数特性を Fig. 5.19 に示す．図中，赤実線はリファレンスとなる目標システムの実測インパルス応答の周波数特性であり，青実線が目標モデルの周波数特性である．図は，実験 1 と同じく応答全体の振幅包絡および位相特性は，おおよそ近似できていることを示している．モデルの次数を 128 次に制限したため，Fig. 5.13 と比較して，振幅特性，位相特性ともに包絡が緩やかになっている．しかし，14 kHz を超える高周波数帯域で歪みが増加する傾向は実験 1 と同様であった．

この歪みを SD を用いて定量化し，制御点ごと，仮想音源ごとの一覧としてまとめたのが Table 5.10 である．表より，仮想音源に対して同側耳側の制御点では歪みが小さく抑えられているが，頭部の陰となる対側耳側の制御点では拡大する傾向が伺える．実験 1 と比較して，歪みはやや大きくなる結果となった．

##### II. 出力追従制御器の設計精度

次に，設計した出力追従制御器と制御対象モデルの直列結合系の周波数特性を Fig. 5.20 に示す．リファレンスは目標モデルの周波数特性とした．図中，赤実線はリファレンスとなる目標モデルの周波数特性であり，青実線は直列結合系の周波数特性である．提案法に

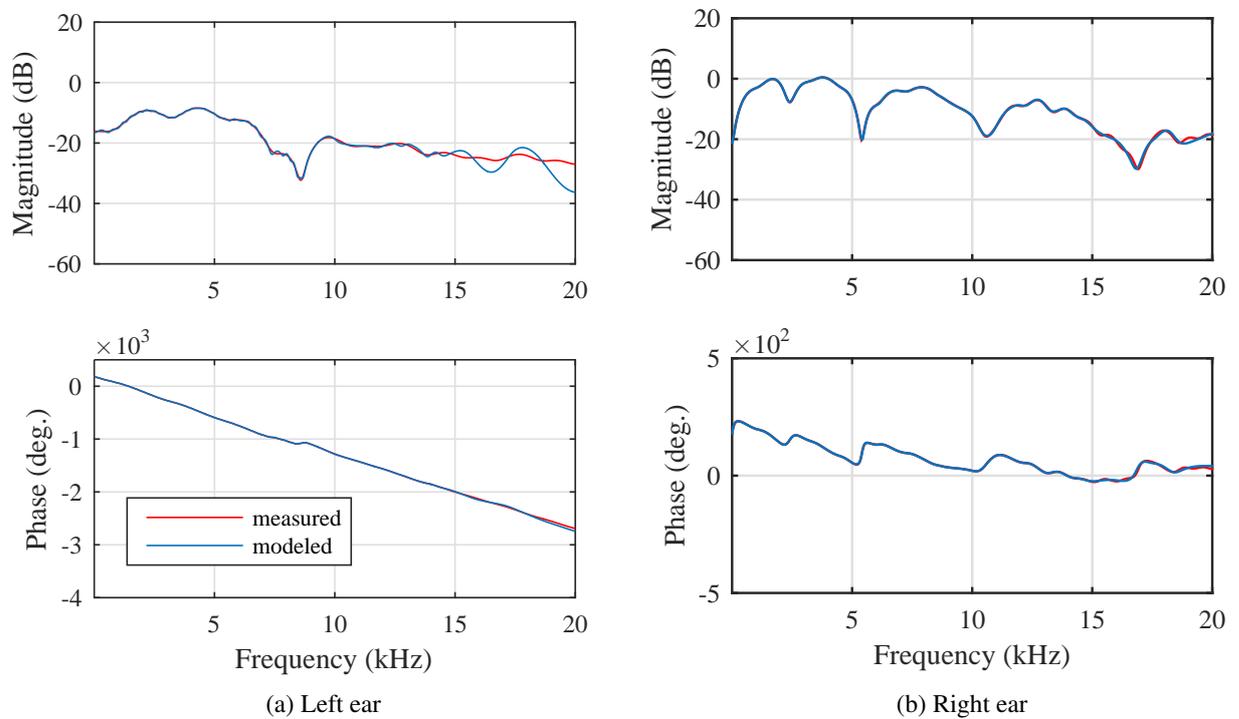


Fig. 5.19: Comparison of frequency characteristics (red solid line: measured, blue solid line: target model)

Table 5.10: SDs generated by modeling error (dB)

Control point	1	2	3	4	5
Left ear	1.34	2.22	2.13	2.42	5.56
Right ear	4.47	1.48	2.76	1.25	1.03

よりほぼ理想的な出力追従制御が可能であることは実験 1 で示されており，本実験でもそれを追認する結果となった．

Table 5.11 に，仮想音源ごとに設計される出力追従制御器の  $\mathcal{H}_\infty$  ノルムを示す．参考として，提案法の Step 2，式 (4.46) の  $\hat{\Phi}$  として Moore-Penrose の擬似逆行列を用いた場合も併記した．Table 5.11 は，提案法により，すべての仮想音源に対して制御器の  $\mathcal{H}_\infty$  ノルムが大幅に小さくなっていることを示している．実験 1 と比較しても， $\mathcal{H}_\infty$  ノルムがおおよそ小さく抑えられる結果となった．これは，モデルの次数を 128 次に制限したことに加え，反射の減少により良条件での求解が可能となったためと考える．また，受聴点が 2 次

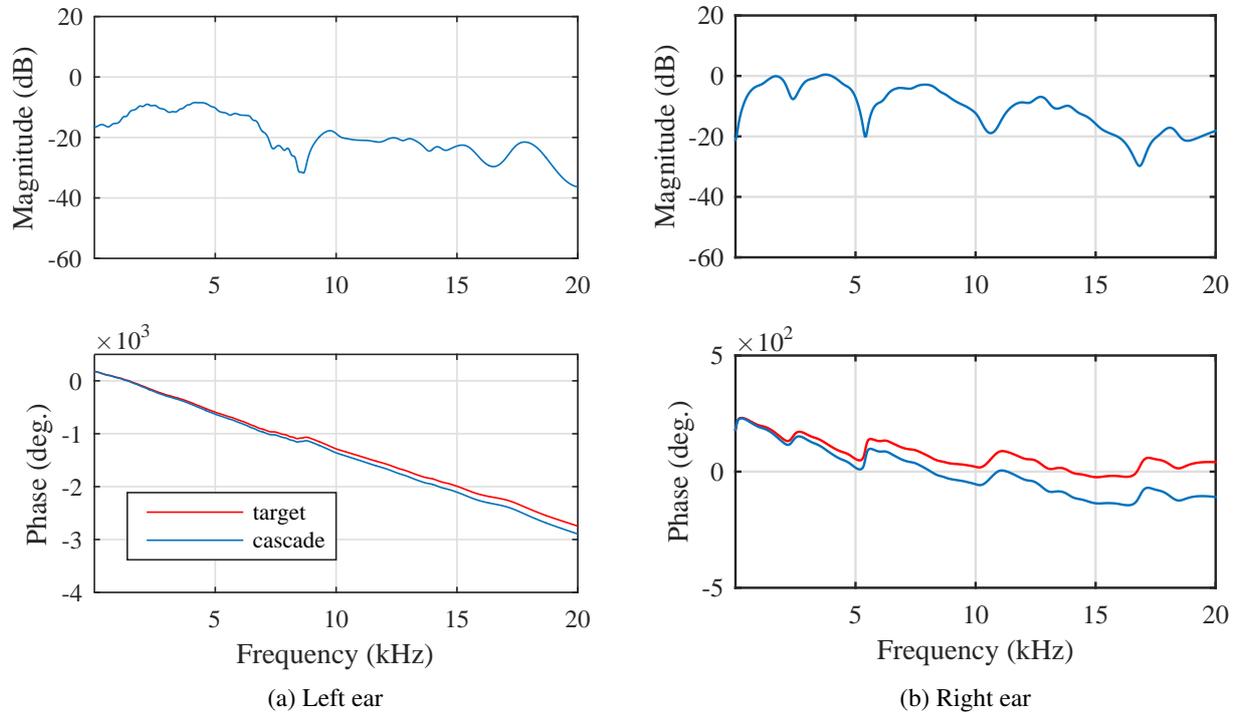


Fig. 5.20: Comparison of frequency characteristics (red solidline: target model, blue solid line: cascade structured system)

Table 5.11:  $\mathcal{H}_\infty$  norm of the tracking controllers

Method	1	2	3	4	5
Proposed	3.46	2.45	0.85	2.80	6.75
Moore-Penrose	71.99	70.98	65.72	28.92	31.35

音源に正対する位置より右方向に移動したときの SD の拡大を，制御点ごと，仮想音源ごとに Table 5.12 にまとめる．受聴点が移動することにより歪みが拡大するが，制御点に対して同側耳側の仮想音源および  $\mathcal{H}_\infty$  ノルムの小さい後方の仮想音源では，ほかの仮想音源と比較して歪みの拡大が小さく抑制されていることが分かる．しかし，移動範囲において同じ音像の定位感を提示するには必ずしも十分抑制性能とはいえない結果となった．

Table 5.12: Increase of SDs (dB)

(a) Left ear					
Distance (cm)	1	2	3	4	5
0.0	0.00	0.00	0.00	0.00	0.00
3.0	4.55	9.21	5.56	11.56	14.08
6.0	6.71	12.60	7.26	15.23	20.14
9.0	5.66	10.62	7.66	15.28	17.39

(b) Right ear					
Distance (cm)	1	2	3	4	5
0.0	0.00	0.00	0.00	0.00	0.00
3.0	12.31	12.78	5.21	7.42	6.23
6.0	14.41	14.88	10.01	11.06	7.11
9.0	14.96	14.70	10.03	12.72	8.31

### III. 出力追従制御器の再生精度

最後に、実際のトランスオーラル再生を想定し、設計した出力追従制御器のインパルス応答に制御対象の実測インパルス応答を畳み込んだ応答の周波数特性を Fig. 5.21 に示す。図中、赤実線はリファレンスとした目標システムの実測インパルス応答から求めた周波数特性であり、青実線が畳み込み出力、すなわち出力追従制御出力の周波数特性である。図は、応答全体の振幅包絡や位相特性、振幅特性上のスペクトルキューなど、音像の定位知覚に係る手掛かりはおおよそ再現されていることを示している。しかし、実験 1 と同じく、目標システムのモデル化誤差を評価した Fig. 5.19 と比較して、明らかに歪みが増加する結果となった。

制御点ごと、仮想音源ごとの SD を Table 5.13 にまとめる。実験 1 とは傾向がやや異なり、仮想音源に対して同側耳側の制御点ほど歪みが減少し、対側耳側の制御点ほど歪みが増加する結果となった。

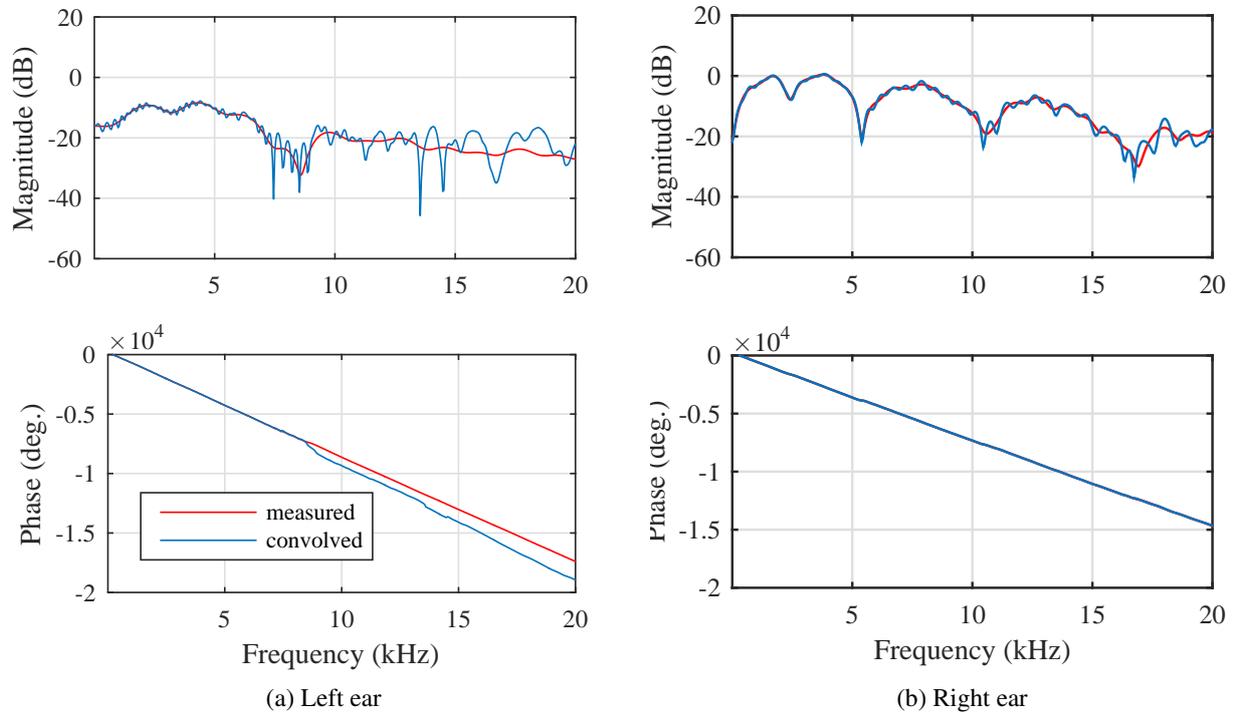


Fig. 5.21: Comparison of frequency characteristics (red solid line: measured, blue solid line: synthesized)

Table 5.13: SDs at control points (dB)

Control point	1	2	3	4	5
Left ear	2.10	3.61	3.64	5.62	7.68
Right ear	5.10	5.19	4.44	3.18	3.00

## 5.5 実験の考察

トランスオーラル再生において、その制御器のロバスト性は、音像の定位感、音色、受聴者の個人性、受聴範囲などに影響を与える重要な要素である。そのロバスト性と深く関係し、再生音場の不確かさや入力信号に含まれる雑音成分を増幅する制御器のゲインを最小化する提案法は、様々な受聴環境を想定した設計アプローチであるといえる。しかし、凸計画問題では閉じた実行可能領域内で最適解を探索するため、必ずしも所望のロバ

スト性を満足する解が得られるとは限らない．実験 2 では受聴点の移動に対する歪みの拡大を評価したが，移動範囲において同じ音像の定位感を提示するのに十分な抑制性能は得ることができなかった．このように厳しい条件のもとで実際的な解を得るためには，従来の正則化法や TSVD 法と同様，制御点での出力追従性能をある程度犠牲にして良条件での求解を可能とする緩和処理が必要となってくる．また，提案法では，制御対象を  $m_p$  入力  $p$  出力 ( $m_p > p$ ) の離散時間線形時不変システムとしている．これは，式 (4.46) により定式化される出力追従制御の制約下で実際的な解を得るためである． $m_p$  に対して  $p$  が増加するに従い良条件での求解が難しくなり， $m_p = p$  の条件下では，通常解を得ることができない．この場合にも，同様の緩和処理が必要となる．

実験 1 ではシステムの次数を 256 次に，実験 2 では 128 次にしたため，仮想音源の位置によってはモデル化誤差がやや大きくなる結果となった．両実験の比較により，次数を高くすればこの誤差が改善されることが分かる．しかし，制御に用いる公称モデルとしては，低い次数で簡単に表現できることが望ましい．また，次数を高くすることにより制御器設計の計算コストが上昇すると共に，再生時の信号処理量も増加する．更には，過剰適合により良条件での求解が難しくなり，結果として制御器のゲインが大きくなってしまうことも考えられる．そのため，これらのトレードオフを考慮したうえで，最適な次数を試行錯誤により決定する必要があると考える．

同時推定と 1 方向ずつの測定の比較では，設計される制御器の精度に大きな差がないことが示された．ただし，既に述べたように，十分な SN 比の確保が同時推定法の課題となっている．SN 比が低下するに従いインパルス応答の無音区間の判定が難しくなり，特にインパルス応答の立ち上がりの判定が困難となる．今後，より多くの 2 次音源を使ってトランスオーラル再生制御器を設計する際には，この SN 比の差が制御器の精度として現れてくることも考えられる．また，このインパルス応答の立ち上がりが最適解を探索する際の実行可能領域を決定するため，適切な測定法の選択には HRTF の測定以上に注意を払う必要がある．

SN 比からはやや論点がずれるが，このインパルス応答の立ち上がりの振幅が大きいほど制御器の  $\mathcal{H}_\infty$  ノルムは小さく抑えられる傾向にある．インパルス応答の先頭を打ち切ればそのモデル化誤差は増加するが，その誤差が聴感上許容される範囲であれば，提案法における緩和処理として用いることができる．その詳細な検討は，今後の重要な課題の一つである．



## 第 6 章

# 結論

本論文では，予測誤差法による HRTF の多方向同時推定法と，出力追従制御を応用したトランスオーラル再生法を提案した．

### 6.1 研究の成果

第 2 章では，本論文で提案するトランスオーラル再生法の基礎となる HRTF について，その定義から測定，応用までを概説した．また，HRTF に係る諸課題を，測定，応用となるバイノーラル再生法，トランスオーラル再生法のそれぞれについて整理した．

第 3 章では，HRTF の測定時間を短縮する方法として，予測誤差法による多方向同時推定法を提案し，その同定入力の一つの  $M$  系列から作成する方法を示した．また，測定を実施する実験室の反射や残響を考慮した同定入力の作成法を提案した．HRTF の多方向同時推定実験を実施し，許容限の範囲となる精度で HRTF を推定することができることを示した．また，実験を実施した防音室の残響時間を考慮して同定入力を作成し，その入力を用いた多方向同時推定実験を実施した．その結果，HRTF の推定精度が，特に頭部の陰となる方向で大幅に改善されることを示した．

第 4 章では，出力追従制御を応用した状態空間でのトランスオーラル再生法を提案した．トランスオーラル再生法では，制御器のゲインにより増幅されるその不確かさや入力信号に混入する雑音成分が問題となる．提案法では，トランスオーラル再生制御器の設計を，制御器の  $\mathcal{H}_\infty$  ノルムを評価指標とする凸計画問題として定式化した．制約条件の記述には LMI を用いた．逆システムの成立する制約条件を出力追従制御の成立する制約条件として緩和し，適当な等価変換と変数返還を施すことにより，トランスオーラル再生制御器の設計に係る制御仕様が LMI に帰着されることを示した．

第 5 章では，提案するトランスオーラル再生法の応用先として研究を進める，22.2ch 音響方式とディスプレイ体型スピーカについて概説した．また，枠型スピーカアレイとラ

インアレイスピーカを用いたトランスオーラル再生をシミュレーション実験により検証し、提案法により制御器のゲインが抑制されること、モデル化精度を高めれば理論的にはほぼ理想的な出力追従制御が可能であること、一方でモデル化精度と制御器のゲインはトレードオフの関係にあることを示した。また、受聴点の移動に対して十分なロバスト性を得るまでには至らなかったこと、そのような厳しい条件のもとで実際的な解を得るためには、出力追従性能をある程度犠牲にして良条件での求解を可能とする緩和処理が必要となってくることを今後の検討課題として示した。

## 6.2 今後の発展性

HRTFの多方向同時推定法では、測定信号を複数のスピーカより同時に再生し、その応答を被測定者の外耳道入口に据え付けられたマイクロフォンで收音する。このとき、被測定者の負担とならないよう音圧レベルは低めに調整されるため、一つ一つの測定方向を考えたとき十分なSN比を確保できていない場合がある。音圧レベルを上げずにSN比を改善するには、測定信号を時間軸方向に伸長する、すなわちシフトレジスタの段数をより多くして作成したM系列を入力信号として使用する方法が考えられる。しかし、測定信号が長くなり、同定入出力ベクトルが大きくなると、予測誤差法の計算過程で必要となる計算時間および計算機上の作業領域が $O(N^3)$ のオーダーで増加する。この問題に対して金井ら[100]は、行列のブロック対角化近似を用いた低計算量推定法を提案した。対角行列のスパース性を利用した計算量削減の有力なアプローチであり、今後の進捗が期待される。

トランスオーラル再生法を含む音場再現・再生法において、その受聴範囲の拡大は、再生法によらず共通する課題の一つである。第5章の実験2では、 $\mathcal{H}_\infty$ ノルムが小さくなるほど受聴者の移動に対して歪の拡大は抑制されるが、制御点のごく近傍を除いて十分な抑制性能とはいえないことが示された。受聴範囲を広く拡大するためには、受聴点を複数とする多点制御に拡張したうえで、提案法を適用していくのが一つのアプローチであると考えられる。ただし、前述のように良条件での求解が難しくなるため、システムの要求仕様に応じて何らかの緩和処理と組み合わせる必要がある。この緩和処理法の検討は、今後の重要な検討課題である。また、人は周波数ごとに様々な手掛かりを用いて音像を定位知覚しており、その手掛かりの重要度は周波数により異なることが知られている。提案法では、全周波数帯域を通じて $\mathcal{H}_\infty$ ノルムを制御器設計の評価指標としたが、想定される再生音場の不確かさや様々な雑音成分の入力に対して高いロバスト性を実現するためには、その性質と共に、人の聴覚特性を考慮した評価指標の検討も必要になってくると考える。著者[101]は、非負値制約に基づく逆システムのゲイン抑制法を提案している。これは、位相がほぼランダム化される高周波数帯域において、同相加算によるゲイン増幅の抑制を設計の制約とするものである。人の聴覚特性を利用した帯域別制御の取り掛かりとして、

今後も検討を継続する所存である。



# 謝辞

本研究は、著者が慶應義塾大学大学院理工学研究科基礎理工学専攻物理情報システム専修後期博士課程在学中に、同大学理工学部物理情報工学科足立修一教授の指導のもとに行ったものです。指導教員として終始親切なご指導とご鞭撻、審査の主査をしていただきました足立修一教授には、謹んで深謝の意を表します。

また、本論文をまとめるにあたり、修正の都度丁寧にご査読いただき、有益なご教示と厳しくも的確なご助言をいただきました副査の慶應義塾大学理工学部システムデザイン工学科大森浩充教授、同システムデザイン工学科滑川徹教授、同物理情報工学科内山孝憲教授には、心からの感謝の意を表します。

本研究は、著者の在学以前から続く慶應義塾大学理工学部物理情報工学科足立研究室とNHK放送技術研究所との共同研究をその礎としています。教職員として従事されました、日本学術振興会特別研究員(PD)丸田一郎博士(現京都大学大学院情報学研究科システム科学専攻助教)、慶應義塾大学理工学部物理情報工学科井上正樹助教をはじめ、共同研究に従事されました、足立研究室OB・OGの竹中裕司氏、徳住ゆりか氏、石川健太郎氏、金井関利氏、菅谷真帆氏、足立研究室M1の森翔平氏、B4の吉村翔氏には、その尽力に心より感謝いたします。特に筆者と在学期間の重なった金井氏、菅谷氏からは、ディスカッションを通して多くの示唆と閃きをいただきました。金井氏とは頭部伝達関数の同時推定法の研究を、菅谷氏とはトランスオーラル再生法の研究をともに取り組み、彼らの力により研究は飛躍的に進捗しました。また、森氏、吉村氏の研究に対する高いモチベーションと勤勉さからは、今も刺激を受けています。

本論文で示した枠型スピーカアレイとラインアレイスピーカの開発は、フォスター電機株式会社およびシャープ株式会社とNHK放送技術研究所との共同研究で進められました。フォスター電機株式会社戦略開発本部森田雄一氏、シャープ株式会社研究開発事業本部岩内謙一室長、同研究開発事業本部服部永雄主事、末永健明主事には、そのご助力に心より感謝いたします。

以上の方々をはじめとして、著者の後期博士課程入学にご配慮いただいた NHK 放送技術研究所の皆様、門外の徒であった著者を暖かく迎え入れていただいた慶應義塾大学工学部物理情報工学科足立研究室の皆様、事務処理ではいつも迷惑をかけてきた秘書の斉藤由美さんには、心より感謝いたします。

最後に、在職中の入学を支援し、仕事と学業との両立のため荒んだ生活を支えてくれた妻稚子に感謝します。

2017年2月

松井 健太郎

## 参考文献

- [1] 安藤彰男, “高臨場感音響技術とその理論,” 電子情報通信学会 基礎・境界ソサイエティ, *Fundamentals Review*, Vol. 3, No. 4, pp. 33–46 (2009)
- [2] M. Camras, “Approach to recreating a sound field,” *Journal of the Acoustical Society of America*, Vol. 43, No. 6, pp. 1425–1438 (1968)
- [3] A. J. Berkhout, D. de Vries, and P. Vogel, “Acoustic control by wave field synthesis,” *Journal of the Acoustical Society of America*, Vol. 93, No. 5, pp. 2764–2778 (1993)
- [4] 西村竜一, “アンビソニックス,” 映像情報メディア学会誌, Vol. 68, No. 8, pp. 616–620 (2014)
- [5] 小野一穂, “マルチチャンネルオーディオ,” 映像情報メディア学会誌, Vol. 68, No. 8, pp. 604–607 (2014)
- [6] 日本オーディオ協会 DHT 普及委員会 WG-2, “サラウンドスピーカの家庭再生配置における許容度調査 (連載 その1),” *JAS Journal* 2012, Vol. 52, No. 1, pp. 55–66 (2012)
- [7] 日本オーディオ協会 DHT 普及委員会 WG-2, “サラウンドスピーカの家庭再生配置における許容度調査 (連載 その2),” *JAS Journal* 2012, Vol. 52, No. 2, pp. 5–22 (2012)
- [8] 永田悠, 立蔵洋介, 猿渡洋, 鹿野清宏, “音場再現システムにおける環境変化に適応的な逆フィルタの逐次的緩和アルゴリズム,” 電子情報通信学会誌 A, Vol. J86-A, No. 8, pp. 824–834 (2003)
- [9] 野田祐樹, 鮫島俊哉, “ゲインスケジューリング制御理論によるアクティブ音響エネルギー制御,” 日本音響学会誌, Vol. 70, No. 2, pp. 53–61 (2014)
- [10] 小野一穂, 松井健太郎, 西口敏行, “マルチチャンネル音響收音マイクロホンの指向性に関する検討,” 電子情報通信学会報告 EA, 応用音響, Vol. 113, No. 503, pp. 31–36 (2014)
- [11] T. Samejima, Y. Sasaki, I. Taniguchi, and H. Kitajima, “Robust transaural sound reproduction system based on feedback control,” *Acoustical Science & Technology*, Vol.

- 31, No. 4, pp. 251-259 (2010)
- [12] B. C. J. Moore, 大串健吾, “聴覚心理学概論,” 誠信書房 (1994)
- [13] J. Blauert, “Spatial Hearing, Revised Edition: The Psychophysics of Human Sound Localization,” The MIT Press (1996)
- [14] J. Blauert, 森本政之, 後藤敏幸, “空間音響,” 鹿島出版会 (1986)
- [15] “新版 音響用語辞典,” 日本音響学会編 (2003)
- [16] 飯田一博, 森本政之, “空間音響学,” コロナ社 (2010)
- [17] K. Iida, M. Itoh, A. Itagaki, and M. Morimoto, “Median plane localization using a parametric model of the head-related transfer function based on spectral cues,” *Applied Acoustics*, Vol. 68, No. 8, pp. 835–850 (2007)
- [18] N. Aoshima, “Computer-generated pulse signal applied for sound measurement,” *Journal of the Acoustical Society of America*, Vol. 69, No. 5, pp. 1484-1488 (1981)
- [19] 浅野太, “Q&A コーナー,” *日本音響学会誌*, Vol. 64, No. 6 p. 370 (2008)
- [20] Y. Suzuki, F. Asano, H. Kim, and T. Sone, “An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses,” *Journal of the Acoustical Society of America*, Vol. 97, No. 2, pp. 1119–1123 (1995)
- [21] 藤本卓也, “低域バンドでの SN 比改善を目的とした TSP 信号に関する検討 -高調波歪の除去-,” *日本音響学会講演論文集*, 3-P-8 (2000)
- [22] 守谷直也, 金田豊, “Logarithmic TSP 信号を用いた高調波歪の検討,” *日本音響学会講演論文集*, 3-P-21 (2004)
- [23] B. B. Bauer, “Stereophonic Earphones and Binaural Loudspeakers,” *Journal of the Audio Engineering Society*, Vol. 9, No. 2, pp. 148–151 (1961)
- [24] B. S. Atal and M. R. Schroeder, “Apparent Sound Source Translator,” US Patent 3,236,949
- [25] M. R. Schroeder, “Digital Simulation of Sound Transmission in Reverberant Spaces,” *Journal of the Acoustical Society of America*, Vol. 47, No. 2, pp. 424–431 (1970)
- [26] M. R. Schroeder, “Models of Hearing,” *Proceedings of the IEEE*, Vol. 63, No. 9, pp. 1332–1350 (1975)
- [27] P. Damaske and V. Mellert, “Ein Verfahren zur richtungstreu Schallabbildung des oberen Halbraumes über zwei Lautsprecher,” *Acustica*, No. 22, pp. 153–162 (1969/70)
- [28] P. Damaske, “Head-related two-channel stereophony with loudspeaker reproduction,” *Journal of the Acoustical Society of America*, No. 50, pp. 1109–1115 (1971)
- [29] D. H. Cooper and J. L. Bauck, “Prospects for Transaural Recording,” *Journal of the Audio Engineering Society*, Vol. 37, No. 1/2, pp. 3–19 (1989)
- [30] J. L. Bauck, “Generalized Transaural Stereo and Applications,” *Journal of the Audio*

- Engineering Society, Vol. 44, No. 9, pp. 683–705 (1996)
- [31] 安藤四一, 設楽貞樹, 前川純一, 城戸健一, “電気計算機による室内音響設計法に関する二, 三の考察,” 日本音響学会誌, Vol. 29, No. 3, pp. 151–159 (1973)
- [32] 平原達也, 大谷真, 戸嶋巖樹, “頭部伝達関数の計測とバイノーラル再生にかかわる諸問題,” 電子情報通信学会 基礎・境界ソサイエティ, Fundamentals Review, Vol. 2, No. 4, pp. 68–85 (2009)
- [33] D. N. Zotkin, R. Duraiswami, E. Grassi, and N. A. Gumerow, “Fast head-related transfer function measurement via reciprocity,” Journal of the Acoustical Society of America, Vol. 120, No. 4, pp. 2202–2215 (2006)
- [34] 松永悟行, 平原達也, “相反法による HRTF 計測の問題点,” 電子情報通信学会技術研究報告 EA, 応用音響, Vol. 109, No. 240, pp. 107–112 (2009)
- [35] 平原達也, “相反法による頭部伝達関数計測に用いる耳栓スピーカの音響特性,” 日本音響学会誌, Vol. 67, No. 8, pp. 331–338 (2011)
- [36] 今井悠貴, 森川大輔, 平原達也, “相反法による頭部伝達関数計測に用いる超小型導電型スピーカユニットの物理特性,” 電子情報通信学会技術研究報告 EA, 応用音響, Vol. 111, No. 270, pp. 67–72 (2011)
- [37] P. Majdak, P. Balazs, and B. Laback, “Multiple exponential sweep method for fast measurement of head-related transfer functions,” Journal of the Audio Engineering Society, Vol. 55, No. 7/8, pp. 623–637 (2007)
- [38] V. Pulkki, M. Laitine, and V. P. Sivonen, “HRTF measurements with a continuously moving loudspeaker and swept sines,” AES 128th Convention, London, UK (2010)
- [39] K. Fukudome, T. Suetsugu, T. Ueshin, R. Idegami, and K. Takeya, “The fast measurement of head related impulse responses for all azimuthal directions using the continuous measurement method with a servo-swiveled chair,” Applied Acoustics, Vol. 68, No. 8, pp. 864–884 (2007)
- [40] 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, Vol. 55, No. 2, pp. 91–99 (1999)
- [41] 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, Vol. 57, No. 11, pp. 685–692 (2005)
- [42] 西野隆典, “頭部伝達関数の方向依存性と個人依存性に対する補間・推定手法,” 日本音響学会誌, Vol. 61, No. 7, pp. 410–415 (2005)
- [43] M. Matsumoto, M. Tohyama, and H. Yanagawa, “A method of interpolating binaural impulse responses for moving sound images,” Acoustical Science & Technology, Vol. 24, No. 5, pp. 284–292 (2003)
- [44] 松本光雄, 柳川博文, “関数近似による水平面内両耳インパルス応答の補間精度の改

- 善,” 日本音響学会誌, Vol. 61, No. 11, pp. 621–628 (2005)
- [45] 渡邊貫治, “頭部伝達関数の補間と合成に関する研究,” 東北大学大学院博士学位論文 (2005)
- [46] A. W. Mills, “Lateralization of high-frequency tones,” *Journal of the Acoustical Society of America*, Vol. 32, No. 132, pp. 132–134 (1960)
- [47] I. J. Hirsh, “Masking of Speech and Auditory Localization,” *International Journal of Auditory*, Vol. 10, No. 2, pp. 110–114 (1971)
- [48] 小澤賢司, 金澤永治, 鈴木陽一, “ヘッドホンを用いたバイノーラル再生における個人性補正の効果,” 日本バーチャルリアリティ学会論文誌, Vol. 5, No. 3, pp. 949–956 (2000)
- [49] 岩谷幸雄, “バイノーラル収録・再生における課題 -ヘッドホンの自由空間等価特性の影響-,” 日本音響学会誌, Vol. 61, No. 7, pp. 386–391 (2005)
- [50] P. A. Nelson, F. Orduña-Bustamante, and D. Engler, “Experiments on a System for the Synthesis of Virtual Acoustic Sources,” *Journal of the Audio Engineering Society*, Vol. 44, No. 11, pp. 990–1007 (1996)
- [51] H. Tokuno, O. Kirkeby, P. A. Nelson, and H. Hamada, “Inverse Filter of Sound Reproduction Systems Using Regularization,” *IEICE Transaction on Fundamentals*, Vol. E80-A, No. 5, pp. 809–820 (1997)
- [52] O. Kirkeby, P. A. Nelson, and H. Hamada, “The ‘Stereo Dipole’: A Virtual Source Imaging System Using Two Closely Spaced Loudspeakers,” *Journal of the Audio Engineering Society*, Vol. 45, No. 5, 387–395 (1998)
- [53] O. Kirkeby, P. A. Nelson, and H. Hamada, “Local sound field reproduction using two closely spaced loudspeakers,” *Journal of the Acoustical Society of America*, Vol. 104, No. 4, pp. 1973–1981 (1998)
- [54] O. Kirkeby and P. A. Nelson, “Digital Filter Design for Inversion Problems in Sound Reproduction,” *Journal of the Audio Engineering Society*, Vol. 47, No. 7/8, pp. 583–595 (1999)
- [55] 浜田晴夫, 得能宏則, O. Kirkeby, P. A. Nelson, “新し音場再生方式 (Stereo Dipole) に関する研究,” 電子情報通信学会技術研究報告 EA, 応用音響, Vol. 96, No. 13, pp. 9–17 (1996)
- [56] T. Takeuchi and P. A. Nelson, “Optimal source distribution for virtual acoustic imaging,” ISVR Technical Report, 288, University of Southampton (2000)
- [57] T. Takeuchi and P. A. Nelson, “Optimal source distribution for binaural synthesis over loudspeakers,” *Journal of the Acoustical Society of America*, Vol. 112, No. 6, pp. 2786–2797 (2002)

- [58] 猿渡洋, “逆フィルタによる音場再生技術,” 映像情報メディア学会誌, Vol. 68, No.8, pp. 612–615 (2014)
- [59] 神沼充伸, 伊勢史郎, 鹿野清宏, “周波数領域における最小ノルム解を利用した多チャンネル音場再現システムにおける逆フィルタの設計,” 日本音響学会誌, Vol. 57, No. 3, pp. 175–183 (2001)
- [60] S. Miyabe, M. Shimada, T. Takatani, H. Saruwatari, and K. Shikano “Multi-Channel Inverse Filtering with Selection and Enhancement of a Loudspeaker for Robust Sound Field Reproduction,” Proceedings of IWAENC 2006, pp. 1-4 (2006)
- [61] N. Kamado, “Sound Field Reproduction Integrating Multi-Point Sound Field Control and Wave Field Synthesis,” Doctoral Dissertation, Department of Information Processing, Graduate School of Information Sciend, Nara Institute of Science and Technology (2012)
- [62] 浅井拓朗, 高以良光, 尾本章, “境界音場制御における逆フィルタの設計指針が音場に与える影響,” 日本音響学会講演論文集, 1-Q-31 (2014)
- [63] 李容子, 伊勢史郎, “正則化パラメータに着目した多チャンネル逆システムの最適化設計法,” 日本音響学会誌, Vol. 69, No. 6, pp. 276–284 (2013)
- [64] 田中章, 大谷尚之, 宮腰政明, “音場再現システムのための正則化逆フィルタの設計,” 電子情報通信学会誌 A, 基礎・境界, Vol. J87-A, No. 11, pp. 1466-1467 (2004)
- [65] H. Kurabayashi, M. Otani, K. Itoh, M. Hashimoto, and M. Kayama, “Development of dynamic transaural reproduction system using non-contact head tracking,” Proceeding of 2nd IEEE Global Conference on Consumer Electronics, pp. 12–16, Makuhari, Japan (2013)
- [66] M. Miyoshi and Y. Kaneda, “Inverse filtering of room acoustics,” IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 36, No. 2, pp. 145–152 (1988)
- [67] 佐藤克昌, 寺田泰宏, 勝俣友紀, 西山清, “J-FHF による 2 入力 2 出力系逆同定を用いたトランスオーラルシステム,” 電子情報通信学会技術研究報告 EA, 応用音響, Vol. 115, No. 169, pp. 35–40 (2015)
- [68] 竹中裕司, 足立修一, “最小二乗法による多入力システム同定のための同定入力の生成法,” 計測自動制御学会論文集, Vol. 47, No. 6, pp. 291–293 (2011)
- [69] 金井関利, 松井健太郎, 中山靖茂, 足立修一, “入力設計の改良による頭部伝達関数の多方向同時推定の高精度化,” 日本音響学会誌, Vol.71, No.7, pp. 316-318 (2015)
- [70] 石川健太郎, 徳住ゆりか, 丸田一郎, 足立修一, 松井健太郎, 中山靖茂, “システム同定理論を用いた頭部伝達関数の 3 次元多方向同時推定,” 日本音響学会誌, Vol. 69, No. 7, pp. 321–330 (2013)
- [71] 足立修一, “MATLAB による制御のための上級システム同定,” 東京電機大学出版局

- (2004)
- [72] 竹中裕司, 足立修一, 松井健太郎, 安藤彰男, “部分空間法による頭部伝達関数の推定,” 第 53 回自動制御連合講演会, pp. 557–560 (2011)
- [73] 西野隆典, “三次元音響表示のための頭部伝達関数の推定法に関する研究,” 名古屋大学博士学位論文 (2003)
- [74] K. Matsui and A. Ando, “Estimation of Individualized Head-Related Transfer Function Based on Principal Component Analysis,” *Acoustical Science & Technology*, Vol. 30, No. 5, pp. 338–347 (2009)
- [75] P. L. Falb and W. A. Wolovich, “Decoupling in the Design and Synthesis of Multivariable Control Systems,” *IEEE Transaction on Automatic Control*, Vol. AC-12, No. 6, pp. 651–659 (1967)
- [76] W. M. Wonham, “Linear Multivariable Control: A Geometric Approach Third Edition,” Springer-Verlag, Berlin (1985)
- [77] 山田功, 木下航, 菊池聡子, “真にプロパーな離散時間系に対する遅れ型逆システムの一設計法,” *日本機械学会論文集 C 編*, Vol. 69, No. 682, pp. 1603–1610 (2003)
- [78] 山田功, 菊池聡子, 戸塚明德, 木下航, “真にプロパーな離散時間系に対する安定な遅れ型逆システムの一設計法 (対角化法の提案),” *日本機械学会論文集 C 編*, Vol. 70, No. 690, pp. 407–413 (2004)
- [79] 浅黄義昭, 渡部慶二, 村松鋭一, 有我祐一, “状態フィードバックと逆システムによる非最小位相系に対する非干渉化の体系的設計法,” *計測自動制御学会論文集*, Vol. 41, No. 3, pp. 234–241 (2005)
- [80] M. Sugaya, K. Matsui, Y. Nakayama, and S. Adachi, “Method of Designing Inverse System for Binaural Reproduction over Loudspeakers by Using Diagonalization Method,” *IEEE Multi-Conference on Systems and Control*, Sydney, Australia, pp. 1032–1037 (2015)
- [81] T. Hiraoka, O. Nishihara, and H. Kumamoto, “Automatic path-tracking controller of a four-wheel steering vehicle,” *Vehicle System Dynamics*, Vol. 47, No. 10, pp. 1205–1227 (2009)
- [82] 辻雅司, “原子力発電プラントの負荷追従運転制御に関する研究,” 北海道大学博士学位論文 (1991)
- [83] D. N. Godbole and S. S. Sastry, “Approximate Decoupling and Asymptotic tracking for MIMO systems,” *Proceedings of the 32nd Conference on Decision and Control*, San Antonio, USA, pp. 2754–2759 (1993)
- [84] 林信男, 黒江康明, 丸橋徹, “外乱オブザーバによる非線形システムの非干渉化制御 - DD ロボットの作業空間における非干渉化 -,” *日本ロボット学会誌*, Vol. 13, No. 3,

- pp. 420–428 (1995)
- [85] 徳田誠, 山本透, 門田良実, “ニューラルネットワークを用いた非干渉化補償器の一設計,” 電気学会論文誌 C, Vol. 125, No. 4, pp. 591–599 (2005)
- [86] 岩崎徹也, “LMI と制御,” 昭晃堂, 東京 (1997)
- [87] 蛭原義雄, “LMI によるシステム制御,” 森出版, 東京 (2012)
- [88] 吉瀬章子, “凸計画問題に対する最適化手法 - 内点法と解析の中心,” システム/制御/情報, Vol. 38, No. 3, pp. 155–160 (1994)
- [89] Y. Shishikui, Y. Fujita, and K. Kubota, “Super Hi-Vision - the star of the show!,” EBU Technical Review, pp. 4–16 (2009)
- [90] 鹿喰善明, “スーパーハイビジョンの研究開発,” NHK 技研 R&D, Vol. 137, pp. 4–9 (2013)
- [91] K. Hamasaki, T. Nishiguchi, R. Okumura, Y. Nakayama, and A. Ando, “A 22.2 Multichannel Sound System for Ultrahigh-Definition TV (UHDTV),” SMPTE Motion Imaging Journal, Vol. 117, No. 3, pp. 40–49 (2008)
- [92] Rec. ITU-R BS.1909-0, “Performance requirements for an advanced multichannel stereophonic sound system for use with or without accompanying picture,” International Telecommunication Union (2012)
- [93] 西口敏行, 小野一穂, 渡辺馨, “8K スーパーハイビジョン音響制作システムの開発と標準化動向,” NHK 技研 R&D, No. 148, pp. 12–21 (2014)
- [94] Rec. ITU-R BS.2051-0, “Advanced sound system for programme production,” International Telecommunication Union (2014)
- [95] ARIB STD-B59, “三次元マルチチャンネル音響方式スタジオ規格,” 電波産業会 (2014)
- [96] 澤谷郁子, “家庭におけるマルチチャンネル音響再生技術,” NHK 技研 R&D, Vol. 128, pp. 11–17 (2011)
- [97] T. Sugimoto, K. Matsui, and H. Okubo, “A Loudspeaker Array Frame Reproducing 22.2 Multichannel Sound for Super Hi-Vision Flat Panel Display,” NAB Proceedings 2012, pp. 16–21 (2012)
- [98] J. Löfberg, “YALMIP: a Toolbox for Modeling and Optimization in MATLAB,” Proceedings of 2004 IEEE International Symposium on Computer Aided Control Systems Design, pp. 284–289 (2004).
- [99] J. F. Sturm, “Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones,” Optimization Methods and Software, Vol. 11, No. 1, pp. 625–653 (1999).
- [100] S. Kanai, M. Sugaya, S. Adachi, and K. Matsui, “Low-Complexity Simultaneous Estimation of Head-Related transfer Functions by Prediction Error Method,” Journal of

the Audio Engineering Society, Vol. 64, No. 11, pp. 895–904 (2016)

- [101] 松井健太郎, “非負値行列因子分解に基づくトランスオーラル再生法の検討,” 日本音響学会講演論文集, 2-P-21 (2014)

## 付録 A

# 特異値分解法によるインパルス応答からの実現

対象として,  $m$  入力  $p$  出力の離散時間線形時不変システムを考える.

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{w}(k) \quad (\text{A.1})$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) + \mathbf{v}(k) \quad (\text{A.2})$$

ただし,  $\mathbf{x}(k)$ ,  $\mathbf{u}(k)$ ,  $\mathbf{y}(k)$  はそれぞれ状態変数ベクトル, 入力ベクトル, 出力ベクトルであり,  $\mathbf{w}(k)$ ,  $\mathbf{v}(k)$  はそれぞれシステム雑音, 観測雑音である.  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$  はシステムの係数行列を表す. また, システムの次数を  $n$  とする.

1. システムのインパルス応答行列  $\mathbf{g}(k)$  を要素として持つブロックハンケル行列  $\mathbf{H}_{i,j}$  を構成する.

$$\mathbf{H}_{i,j} = \begin{bmatrix} \mathbf{g}(1) & \mathbf{g}(2) & \cdots & \mathbf{g}(j) \\ \mathbf{g}(2) & \mathbf{g}(3) & \cdots & \mathbf{g}(j+1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{g}(i) & \mathbf{g}(i+1) & \cdots & \mathbf{g}(i+j+1) \end{bmatrix} \quad (\text{A.3})$$

ただし,  $i, j$  はシステムの次数  $n$  より十分に大きいものとする.

2.  $\mathbf{H}_{i,j}$  を特異値分解する.

$$\begin{aligned} \mathbf{H}_{i,j} &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\top} \\ &= \begin{bmatrix} \mathbf{U}_s & \mathbf{U}_w \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_w \end{bmatrix} \begin{bmatrix} \mathbf{V}_s^{\top} \\ \mathbf{V}_w^{\top} \end{bmatrix} \end{aligned} \quad (\text{A.4})$$

ただし, 下付き文字 's' は信号部分空間, 'w' は雑音部分空間を意味する. また,  $\mathbf{\Sigma}_s$  は  $\mathbf{H}_{i,j}$  の特異値を大きいものから対角に並べた対角行列であり,  $\mathbf{\Sigma}_w$  は残りの小さな特異値を対角に並べた対角行列である.

$$\mathbf{\Sigma}_s = \text{diag}[\sigma_1, \sigma_2, \cdots, \sigma_n], \quad \sigma_1 > \sigma_2 > \cdots > \sigma_n \quad (\text{A.5})$$

$$\Sigma_w = \text{diag} [\sigma_{n+1}, \sigma_{n+2}, \dots], \quad \sigma_n \gg \sigma_{n+1} > \sigma_{n+2} > \dots \quad (\text{A.6})$$

このとき,  $\Sigma_s$  に含まれる特異値の数がシステムの次数の推定値  $n$  となる.

3. システムのインパルス応答行列は

$$g(k) = \begin{cases} \mathbf{0}, & k < 0 \\ \mathbf{D}, & k = 0 \\ \mathbf{CA}^{k-1}\mathbf{B}, & k > 0 \end{cases} \quad (\text{A.7})$$

で与えられるため,

$$\mathbf{H}_{i,j} = \begin{bmatrix} \mathbf{CB} & \mathbf{CAB} & \dots & \mathbf{CA}^{j-1}\mathbf{B} \\ \mathbf{CAB} & \mathbf{CA}^2\mathbf{B} & \dots & \mathbf{CA}^j\mathbf{B} \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{CA}^{i-1}\mathbf{B} & \mathbf{CA}^i\mathbf{B} & \dots & \mathbf{CA}^{i+j-2}\mathbf{B} \end{bmatrix} \quad (\text{A.8})$$

となる. これを念頭におき,  $\mathbf{H}_{i,j}$  を

$$\begin{aligned} \mathbf{H}_{i,j} &\approx \mathbf{U}_s \Sigma_s \mathbf{V}_s^\top \\ &= \mathbf{\Gamma}_{1:i} \mathbf{\Omega}_{1:j} \end{aligned} \quad (\text{A.9})$$

のように分解する. ただし,

$$\begin{aligned} \mathbf{\Gamma}_{1,j} &= \mathbf{U}_s \Sigma_s^{\frac{1}{2}} \\ &= \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix} \end{aligned} \quad (\text{A.10})$$

$$\begin{aligned} \mathbf{\Omega}_{1:i} &= \Sigma_s^{\frac{1}{2}} \mathbf{V}_s^\top \\ &= \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \dots & \mathbf{A}^{j-1}\mathbf{B} \end{bmatrix} \end{aligned} \quad (\text{A.11})$$

はそれぞれ拡大可観測行列, 拡大可制御行列である.

4. この分解を行うことができれば, 拡大可観測行列と拡大可制御行列の第 1 ブロック要素からそれぞれ  $\mathbf{C}$  と  $\mathbf{B}$  を求めることができる.

$$\mathbf{C} = \mathbf{\Gamma}_{1,j}(1 \dots p, 1 \dots n) \quad (\text{A.12})$$

$$\mathbf{B} = \mathbf{\Omega}_{1:i}(1 \dots n, 1 \dots m) \quad (\text{A.13})$$

なお,  $\mathbf{D} = g(0)$  である.

## 5. 拡大可観測行列は

$$\begin{aligned}
\mathbf{\Gamma}_{2:i} &= \begin{bmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^i \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix} \mathbf{A} \\
&= \mathbf{\Gamma}_{1:i} \mathbf{A}
\end{aligned} \tag{A.14}$$

と表すことができるため， $\mathbf{A}$  の最小 2 乗解は

$$\begin{aligned}
\mathbf{A} &= \mathbf{\Gamma}_{1:i}^+ \mathbf{\Gamma}_{2:i} \\
&= (\mathbf{\Gamma}_{1:i}^\top \mathbf{\Gamma}_{1:i})^{-1} \mathbf{\Gamma}_{1:i}^\top \mathbf{\Gamma}_{2:i}
\end{aligned} \tag{A.15}$$

により与えられる．ただし， $(\cdot)^+$  は  $(\cdot)$  の Moore-Penrose の擬似逆行列を表す．



## 付録 B

# 研究業績

### 学位請求に用いる定期刊行誌掲載原著論文

1. Kentaro Matsui, Satoshi Oishi, Takehiro Sugimoto, Satoshi Oode, Yasushige Nakayama, Hiroyuki Okubo, Hiroshi Sato, Koji Mizuno, Yuichi Morita, and Shuichi Adachi, “Binaural Reproduction of 22.2 Multichannel Sound with Flat Panel Display-Integrated Loudspeaker Frame for Home Use,” 映像情報メディア学会誌, Vol. 68, No. 10, pp. J447–J456 (2014)
2. 松井健太郎, 大石 諭, 杉本岳大, 森田雄一, 足立修一, “22.2 マルチチャンネル音響の枠型スピーカによるバイノーラル再生法の開発,” 映像情報メディア学会誌, Vol. 70, No. 1, pp. J37–J46 (2016)
3. 松井健太郎, 菅谷真帆, 足立修一, “出力追従制御を応用したトランスオーラル再生制御器の設計,” 日本音響学会誌 (掲載決定)

### 学位請求に用いる国際会議発表論文 (査読付き, 発表者\*)

1. Kentaro Matsui\*, Yasushige Nakayama, Maho Sugaya, and Shuichi Adachi, “Binaural Reproduction over Loudspeakers Using Low-Order Modeled HRTFs,” 137th AES Convention, Los-Angeles, USA (2014)
2. Kentaro Matsui\*, Maho Sugaya, and Shuichi Adachi, “Design Method for Transaural Reproduction Controller Based on H-Infinity Norm Constraint,” SICE Annual Conference 2016, Tsukuba, Japan (2016)

## 関連する定期刊行誌掲載原著論文

1. 石川健太郎, 徳住ゆりか, 丸田一郎, 足立修一, 松井健太郎, 安藤彰男, “システム同定理論を用いた頭部伝達関数の 3 次元空間多方向同時推定,” 日本音響学会誌, Vol. 69, No. 7, pp. 321–330 (2013)
2. 金井関利, 松井健太郎, 中山靖茂, 足立修一, “入力設計の改良による頭部伝達関数の多方向同時推定の高精度化,” 日本音響学会誌, Vol. 71, No. 7, pp. 316–318 (2015)
3. Maho Sugaya, Kentaro Matsui, Yasushige Nakayama, and Shuichi Adachi, “Low-Order Modeling of Head-Related Transfer Function for Binaural Reproduction,” SICE Journal of Control, Measurement, and System Integration, Vol. 8, No. 2, pp. 108–113 (2015)
4. Sekitoshi Kanai, Maho Sugaya, Shuichi Adachi, and Kentaro Matsui, “Low-Complexity Simultaneous Estimation of Head-Related Transfer Functions by Prediction Error Method,” Journal of the Audio Engineering Society, Vol. 64, No. 11, pp. 895–904 (2016)

## 関連する国際会議発表 (査読付き, 発表者\*)

1. Maho Sugaya\*, Shuichi Adachi, Kentaro Matsui, and Yasushige Nakayama, “Low-order Modeling of Head-Related Transfer Functions for Transaural Systems,” SICE Annual Conference 2014, Sapporo, Japan (2014)
2. Sekitoshi Kanai\*, Kentaro Matsui, and Shuichi Adachi, “Identification Input Design for Simultaneous Estimation of Head-Related Transfer Functions,” SICE Annual Conference 2014, Sapporo, Japan (2014)
3. Sekitoshi Kanai\*, Kentaro Matsui, Yasushige Nakayama, and Shuichi Adachi, “Uncorrelated Input Signals Design and Identification with Low-Complexity for Simultaneous Estimation of Head-Related Transfer Functions,” 137th AES Convention, Los Angeles, USA (2014)
4. Maho Sugaya\*, Kentaro Matsui, Yasushige Nakayama, and Shuichi Adachi, “Method of Designing Inverse System for Binaural Reproduction over Loudspeakers by Using Diagonalization Method,” IEEE Multi-Conference on Systems and Control, Sydney, Australia (2015)

## 関連する国内会議発表 (発表者\*)

1. 松井健太郎\*, 中山靖茂, 菅谷真帆, 足立修一, “低次 HRTF モデルを用いたトランスオーラル再生,” 電子情報通信学会技術研究報告, EA 応用音響, Vol. 114, No. 178, pp. 25-30 (2014)
2. 松井健太郎\*, 中山靖茂, 足立修一, “非負値行列因子分解のトランスオーラル再生への応用の一検討,” 日本音響学会講演論文集, 2-P-21 (2014)
3. 金井関利\*, 松井健太郎, 中山靖茂, 足立修一, “頭部伝達関数の同時推定のための入力信号設計,” 日本音響学会講演論文集, 2-P-20 (2014)
4. 松井健太郎\*, 金井関利, 菅谷真帆, 足立修一, “頭部伝達関数の共通極・零モデリング法の一検討,” 日本音響学会聴覚研究会資料, Vol. 45, No. 6, pp. 481-484 (2015)
5. 金井関利\*, 松井健太郎, 中山靖茂, 足立修一, “多方向頭部伝達関数の ARX モデリング,” 電子情報通信学会技術研究報告, EA 応用音響, Vol. 114, No. 473, pp. 291-294 (2015)
6. 菅谷真帆\*, 松井健太郎, 中山靖茂, 足立修一, “トランスオーラル再生のための対角化法による逆システムの構成,” 電子情報通信学会技術研究報告, EA 応用音響, Vol. 114, No. 473, pp. 285-290 (2015)

## 関連する NHK 放送技術研究所論文

1. 松井健太郎, “家庭用 22.2 マルチチャンネル音響再生システム,” NHK 技研 R&D, No. 148, pp. 45-54 (2014)
2. Kentaro Matsui, “22.2 Multichannel Sound Reproduction System for Home Use,” Broadcast Technology, No. 59, pp. 10-17 (2015)

## 受賞

1. 2014 年度映像情報メディア学会映像情報メディア未来賞フロンティア賞, 22.2 マルチチャンネル音響の枠型スピーカによるバイノーラル再生法の開発 (2015)