

平成 16 年度

修士学位論文

クロストーク成分の相互相関に着目した 音場再生システム

Sound Field Reproduction System Focused on
Crosscorrelation of Cross-Talk Components

1075059 浜崎 真二

指導教員 福本 昌弘

2005 年 3 月 11 日

高知工科大学大学院 工学研究科 基盤工学専攻
情報システム工学コース

要 旨

クロストーク成分の相互相関に着目した 音場再生システム

浜崎 真二

トランスオーラルシステムに代表されるラウドスピーカ再生による音場再生システムは、受聴者にヘッドホン等を装着させることなく音を提示することができる。しかし受聴点（制御点）において音圧を所望信号に近似させるためにはラウドスピーカ（2次音源）から制御点までの伝達特性の影響を考慮する必要がある。このことは空間伝達特性の逆特性を適応フィルタにより推定し、所望信号に畳み込むことで解決される。現在までに逆特性を近似する手法についてはシステム論的立場から様々な検討がなされてきた。その中で最も有効とされている手法に MINT 理論に基づいた多チャンネル-多点制御系がある。これは制御点数 M に対し 2 次音源数を $M + 1$ 用意することで、逆特性の推定と同時にクロストーク現象を除去し、音の再現を可能にしている。しかし余分なスピーカを用いることで制御系が複雑になり、残響時間が長い場合には演算量を増加させてしまうことが懸念され実用化に至っていない。そこで本論文では制御点数 2 に対し 2 次音源数 2 で構成される新たな制御系を提案している。通常、2 チャンネル-2 点制御系では厳密に伝達特性の逆特性を近似させることは、システム論的見解からは困難とされているが、クロストーク成分に強い相関があるという事実を利用すると、クロストークの影響は軽減可能であり、その結果、原音に対する再現音の再現精度が向上することを計算機シミュレーションにより確認している。

キーワード 音場再生, マルチチャンネル, 相互相関, クロストーク, 聴覚特性

Abstract

Sound Field Reproduction System Focused on Crosscorrelation of Cross-Talk Components

Shinji Hamasaki

The multi-channel sound field reproduction system is one of the most promising methods which can construct the similar acoustic field as the recording area by controlling several points. However, these systems have not been applied to the realization of the acoustic field because of the insufficient knowledge of the multiple channel system's behavior. A loudspeaker should be dealt with multi-path by one controller system because it is able to have only one correct filter. In the past research, sound field reproduction system controls two channels by one correct filter. In this research, the simple multiple sound field reproduction system is constructed by parallel constituting past research's system. It has consisted of two input points and two control points which correspond to both ears of a listener in a room. The result, examining about the transfer characteristic relation of input signals and observation signals, cross-correlation tend to 1 on cross-talk components. Therefore, the parameter of correct filter is decided by the cross-correlation and the reproduction of sound. Moreover, we have showed the tendency to reproduce sound field by computer simulation.

key words sound field reproduction, muti-channel, crosscorrelation, cross-talk, acoustic characteristics

目次

第 1 章	序論	1
1.1	はじめに	1
1.2	従来手法の問題点	5
1.3	研究の目的と本論文の構成	7
第 2 章	音場再生システムと逆特性の近似	8
2.1	はじめに	8
2.2	音響システムと線形モデル	8
2.3	逆特性の近似	10
2.4	多チャンネル音場再生システムにおける従来の逆特性近似	12
2.4.1	多チャンネル-多点制御近似逆フィルタの設計方針	12
2.4.2	MINT 理論に基づく近似逆フィルタ設計	14
2.5	従来法の問題点	16
2.6	まとめ	17
第 3 章	2 チャンネル-2 点制御による多入力信号補正法	18
3.1	はじめに	18
3.2	2 チャンネル-2 点制御系における問題点	18
3.3	多入力信号補正システム	20
3.3.1	逆特性の近似的推定法	20
3.3.2	クロストークの影響を考慮した補正フィルタ係数の導出	23
3.3.3	補正フィルタ係数の更新手順	25
3.4	計算機シミュレーションによる評価	26
3.4.1	室内伝達関数の測定	26

目次

3.4.2	シミュレーション条件	26
3.4.3	シミュレーション結果	28
3.5	まとめ	29
第 4 章	聴覚特性に基づく出力誤差の算出	38
4.1	はじめに	38
4.2	聴覚特性に基づく出力誤差の算出	38
4.2.1	重み付け関数の算出	39
4.2.2	重み付け誤差の算出とフィルタ係数の更新手順	40
4.3	計算機シミュレーションによる評価	42
4.3.1	シミュレーション条件	42
4.3.2	シミュレーション結果	43
4.4	まとめ	43
第 5 章	結言	45
5.1	本研究のまとめ	45
5.2	今後の課題	45
	謝辞	47
	参考文献	49

目次

1.1	マルチスピーカ方式の一例	2
1.2	バイノーラルシステムの一例	3
1.3	トランスオーラルシステムの一例	4
1.4	MINT 法に基づくトランスオーラルシステム	5
1.5	適応フィルタを用いたトランスオーラルシステム	6
2.1	反射音を有する室内における音の伝搬の様子	9
2.2	図 2.1 の等価回路	10
2.3	室内伝達特性の補正	11
2.4	MINT 理論に基づく音場制御系	12
2.5	2 次音源数 I , 制御点数 J の音響系	13
3.1	2 チャネル-2 点制御系線形モデル	19
3.2	2 チャネル-2 点制御系線形モデル (クロストーク無し)	20
3.3	近似逆フィルタと構成図	22
3.4	2 チャネル音場再生システムの構成図	24
3.5	室内インパルス応答の測定環境 1	27
3.6	室内インパルス応答の測定環境 2	27
3.7	室内インパルス応答の測定環境 3	28
3.8	環境 1 の室内インパルス応答 g_{11,L_g}	30
3.9	環境 1 の室内インパルス応答 g_{12,L_g}	30
3.10	環境 1 の室内インパルス応答 g_{21,L_g}	31
3.11	環境 1 の室内インパルス応答 g_{22,L_g}	31
3.12	右側原信号	32

目次

3.13	左側原信号	32
3.14	改善量の算出	33
3.15	右側改善量に対する収束特性	33
3.16	左側改善量に対する収束特性	34
3.17	環境 1 に対する収束特性の比較 (左右改善量の平均)	34
3.18	環境 2 に対する収束特性の比較	35
3.19	環境 3 に対する収束特性の比較	35
4.1	A 特性音圧レベル	39
4.2	重み付け関数	40
4.3	重み付け関数	41
4.4	右チャンネルの収束特性の比較	43
4.5	左チャンネルの収束特性の比較	44

表目次

3.1	インパルス応答の測定条件	27
3.2	環境 1 の改善量 ($r_N = 0.73$)	36
3.3	環境 2 の改善量 ($r_N = 0.55$)	36
3.4	環境 3 の改善量 ($r_N = 0.45$)	37

第 1 章

序論

1.1 はじめに

コンサートホール，映画館，競技場等に足を運ぶと，自分の目でとらえる映像にマッチして前後左右へと独特な音の広がりを感じる．そういった音響空間が仮に自分の部屋にいなから与えられたとしたら，受聴者はまさに仮想世界の中にいるような感じをもつはずである．競技場等で目に入る映像には動きがあり，“臨場感”を創出するためには映像の動きに同期した音像定位や音像移動が不可欠となる．近年，その臨場感を家庭で実現することを目的とした「ホームシアター」が急速に普及している．ホームシアターでは，DVD に記録された映像を大型スクリーンに投影し，音声情報を 6 つのラウドスピーカにより再生する．これらのスピーカは受聴者を取り囲むように配置され，5 つのラウドスピーカと重低音再生専用のスーパーウーファから構成され，5.1ch サラウンドシステムと呼ばれている．従来の左右 2 つのスピーカによるステレオ再生系では受聴者前方の左右への音の拡がり感については，ある程度創出することができた．しかしながら，受聴者の側方や後方，またはスピーカよりも外側への音像定位は不可能であった．これに対し，5.1ch 再生系ではスクリーン中央部にセンタースピーカ，その両サイドに 2 つのメインスピーカ，受聴者後方に 2 つのリアスピーカを配置するため，受聴者側方から後方に至るまで極めて良好に音像を定位させることができるようになった．これから本格始動されようとしている地上デジタル放送では，MPEG 符号化による 5.1ch 対応の音声再生放送も行われ [1]，音の臨場感への要求はますます高まってくると思われる．

音の臨場感を与える技術には大きく 2 つの分類がある．1 つは映像に同期して音像移動を

1.1 はじめに

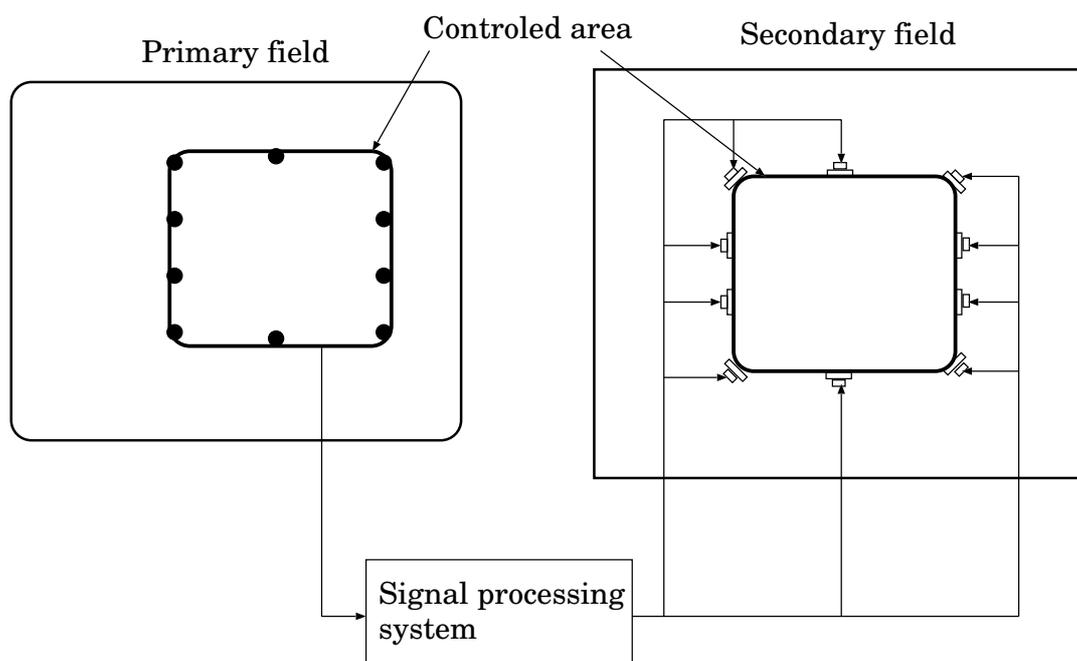


図 1.1 マルチスピーカ方式の一例

実現するものであり、もう 1 つはコンサートホール等の音場を忠実に再現することである。前者については、原音場の特性を物理的に再現することにより、受聴者がそれらしく感じるように再生音場の特性と合成させるものである。これは主に映像に合わせた音環境の合成を目標とした技術である。音像定位を実現するためには、レベル差や時間差、位相差をコントロールすることが必要となる。上述のような 5.1ch 再生方式はこの中に分類される。

一方、原音場の物理的特性を忠実に再現することを目指すためには、再生音場の反射や残響といった特性を考慮して再現する必要がある。この手法の最も基礎的な考え方として、図 1.1 に示すようなマルチスピーカ方式があげられる。これは原音場において複数のマイクロホンを用いて音を収録し、再現音場にてマイクロホンと位置や方向を同一に配置されたスピーカを用いて音を再生する手法である。しかしながら、この方法は多くの技術的困難が伴う。例えば再生する室内は無響室、もしくはきわめて残響の少ない部屋でのみ実現可能であることや、再現領域に広さを求める場合、それに応じた多くのスピーカやマイクロホンが必要となる。以上のような理由から、この方式は実現が困難であると考えられてきた。

マルチスピーカ方式のように波動論的立場から音場再現を試みる一方で、システム論的ア

1.1 はじめに

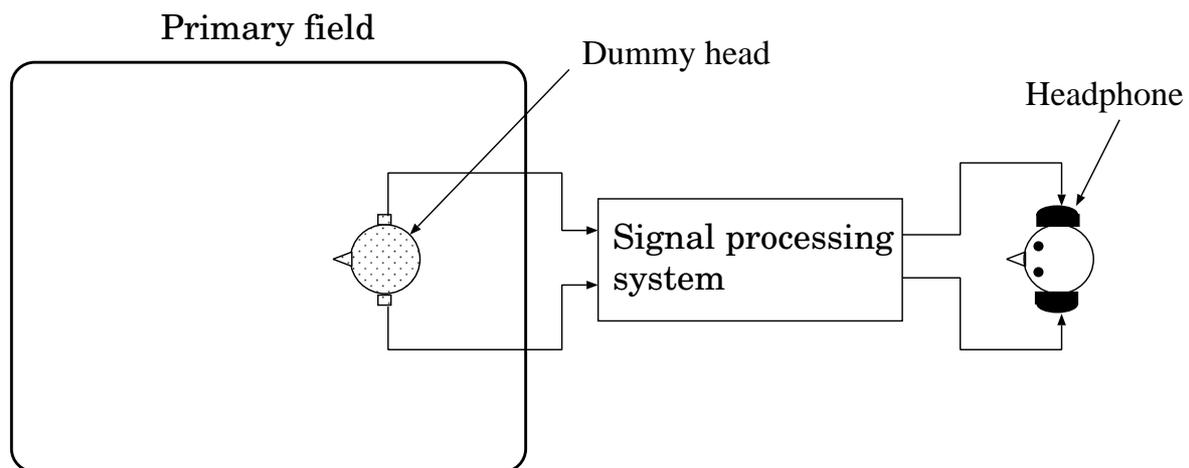


図 1.2 バイノーラルシステムの一例

ブローチから音場再生を達成する手法が考案された。これは、再生する領域を受聴者の耳元付近と限定することにより、音場再生を達成することを目指したものである。このシステムの本質は、受聴者の耳元で再生される信号が原音場で收音時に得られる耳入力信号と一致するように再生することである。これを実現するために原信号は頭部、あるいは擬似頭（ダミーヘッド）を用いて收音される。この收音方法はダミーヘッド收音、またはバイノーラル收音と呼ばれている。

バイノーラル收音した信号を再生するデバイスとしてはヘッドホンの使用が考えられる。これは、ダミーヘッドで收音した信号が受聴者の両耳それぞれに再生されれば、原音場で受聴しているのと同等の聴感が得られるであろうという考えに基づく [2]。ここではこの再生方法をバイノーラル再生、もしくはバイノーラルシステムと呼ぶ。図 1.2 にバイノーラルシステムを示す。バイノーラルシステムはハードウェア構成が簡素であるため、バイノーラル收音された信号さえあれば既存のハードウェアで容易に実現可能となる。しかしながらバイノーラルシステムはヘッドホン再生に起因した問題点が存在する。まず 1 つ目には、ヘッドホン再生に伴う頭内定位の問題である。これはヘッドホン特性の混入、外耳道共振系を 2 重に経ることなどが原因とされている。2 つ目に、受聴者とダミーヘッドの頭部形状の違いによる頭部伝達関数の個人性から起こる定位精度の劣化が挙げられる。3 つ目に、頭部移動

1.1 はじめに

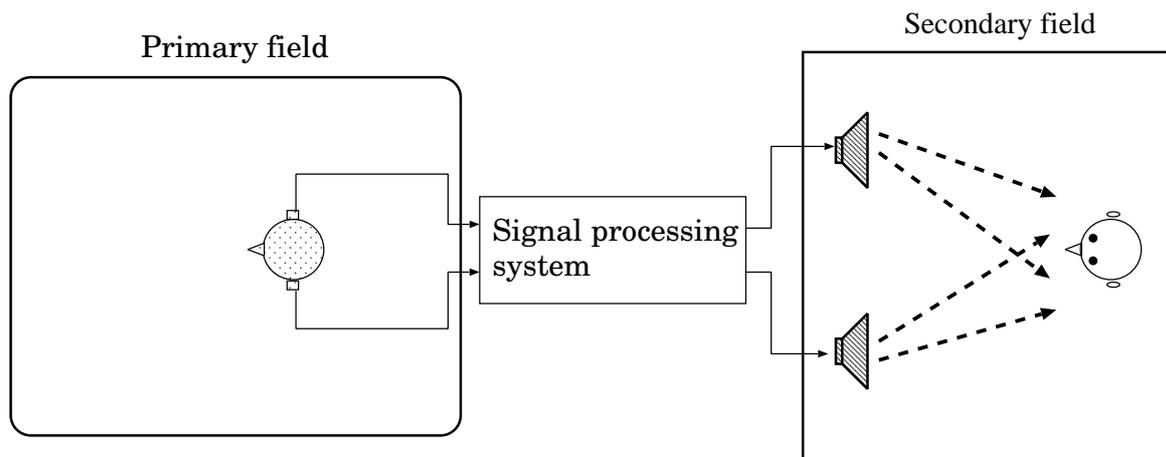


図 1.3 トランスオーラルシステムの一例

(回転)に対する追従性の問題がある．そして4つ目にヘッドホンを直接身に着けることによって受聴者に装着負荷を与えてしまう問題がある．

バイノーラル収録された信号を，ラウドスピーカを用いることにより再生する手法についても数多く検討される．空間に伝達させることによってバイノーラル信号を再現することを目指すため，この再生手法はトランスオーラルと呼ばれる．図 1.3 にトランスオーラルシステムを示す．トランスオーラルシステムにおいては，ラウドスピーカと受聴者耳元（受音点）間の伝達系におけるクロストークを除去し，所望の入力信号をそのまま各受音点のみに再現することが必要となる．クロストークの除去についてはクロストークキャンセラや，OSS (orthostereophonic system) などが提案されてきた．しかしながらトランスオーラルシステムにおいても，再生する室内は無響室や残響の極めて少ない室内に限定されていた．残響が無視できない室内でトランスオーラルシステムを構築するには，MINT 理論 (multi-input/output inverse theorem) に基づき，図 1.4 のように受音点数以上のスピーカを用いれば制御可能であることが知られている．

ラウドスピーカ再生による音場再生システムの構築に際しては，再現音の劣化を防ぐために室内伝達系の変動を考慮しなければならない．室内伝達系の変動要因は大きく分けて2つある．1つ目に物体の移動による伝達経路のよるものがあり，2つ目に温度や湿度，風の影響などの自然環境変化によるものである．この問題を解決する手法に，適応フィルタの使用

1.2 従来手法の問題点

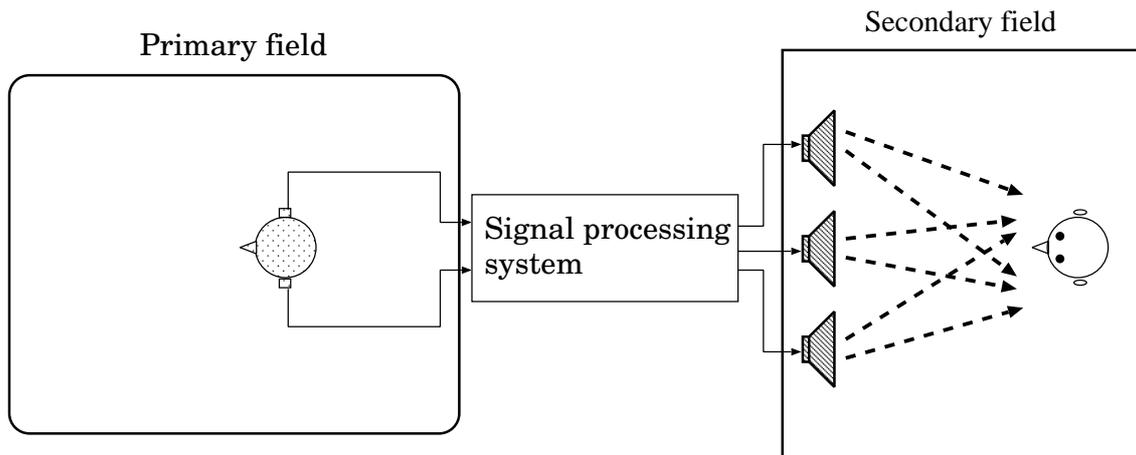


図 1.4 MINT 法に基づくトランスオーラルシステム

がある．適応フィルタとは，出力を目的信号に近づけていく機能を持ったフィルタであり，その多くが最小 2 乗原理に基づいて動作する．図 1.5 は適応処理を行う音場再生システムの例である．この図に示したとおり，全ての制御点にセンサを配置できないようなシステムでは，仮想マイクロホンによる適応システムが提案されている．これらは制御できる帯域が比較的 low 域であるため，能動騒音制御のようなシステムには適しているが，音場再生のように人間の可聴域そのものを扱うことが望ましいようなシステムには適用が困難である．

1.2 従来手法の問題点

以上述べたように，ラウドスピーカ再生による音場再生を達成するためには，ラウドスピーカ-受音点間の反射や残響の影響を除去することが必要となる．なぜならば，反射音の存在する室内においては，再生信号に含まれる反射音と再生音場固有の反射音との相互作用により，所望の音場とは全く異なった性質の過剰な反射音が生産されるからである．よって室内伝達特性を持つような逆フィルタの設計が不可欠となる．

逆フィルタの設計方法については，従来より多くの手法が提案されてきた．ラウドスピーカ数と受音点数が増加するにつれて，逆フィルタ計算に要するメモリ量と演算量は大幅に増加する．また残響を有する室内においては，その影響を除去するための逆フィルタのタップ

1.2 従来手法の問題点

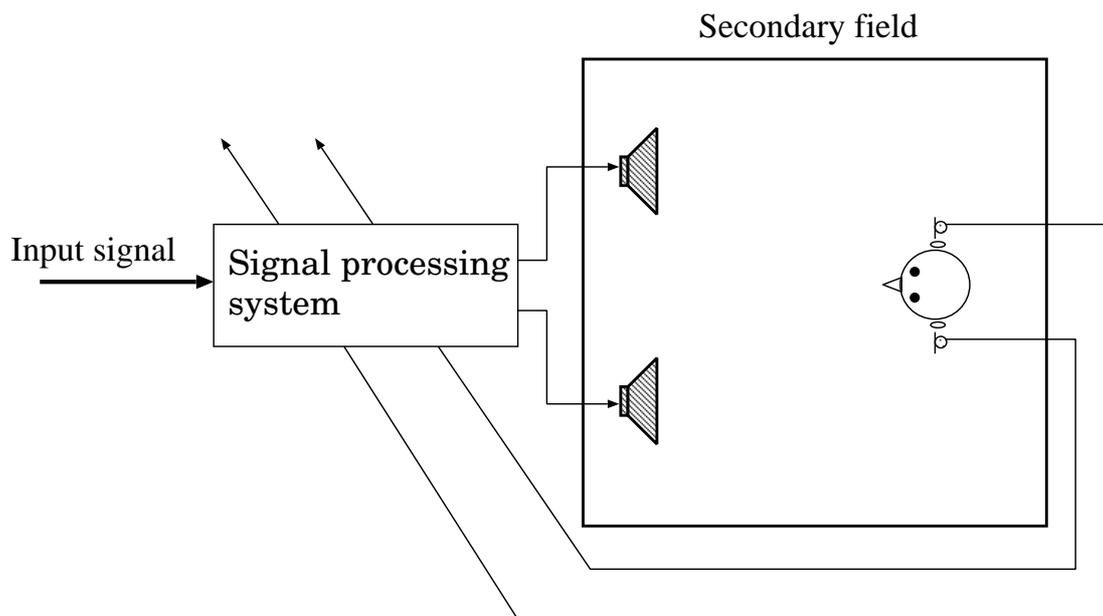


図 1.5 適応フィルタを用いたトランスオーラルシステム

長が非常に長くなることが予想される。しかも、現実の室内においては厳密な逆フィルタを設計できる保障はなく、近似逆フィルタが設計されるが、近似による誤差の影響を考慮する必要がある。

さらに室内の状態は時不変ではなく、温度や湿度、風の影響によってその伝達特性は容易に変動しうる。逆フィルタの設計に用いた室内伝達特性と音場再生を行うときの室内伝達特性が異なった場合、逆フィルタは室内伝達特性の逆特性を持ち得ないため、受音点において再現音の劣化が避けられない。高周波数帯域の音は波長が短いために伝達系の変動の影響を受けやすく、高域になるにしたがって厳密な再現は困難となる。そのため、従来の音場再生システムの評価は上限周波数を 5kHz 程度としている場合が多い。

音場再生システムにおいて再生帯域が狭いことは 2 つの理由により好ましくない。1 つには、人間の可聴域の上限が約 20kHz とされていることから、高域を制限することによって受聴者は提示音をこもったように知覚することが考えられるからである。音場再生を行う際に高域の音を除去することによって、再現音がただしく知覚されないことが想像される。もう 1 つの理由として、人間が音像方向を知覚する際には両耳間レベル差や両耳間位相差だ

1.3 研究の目的と本論文の構成

けでなく、音のスペクトル変化も影響するからである。

これらの理由により、高品質な音場再生を達成するためには、伝達系の変動に対する逆フィルタの追従は不可欠となる。

1.3 研究の目的と本論文の構成

本論文では、前説で述べた逆フィルタの計算量増加に関する問題を解決するために、2チャンネルの音場再生を前提に、適応フィルタを用いた新たな制御システムを提案する。また、伝達系の変動に対する対策として、近似逆フィルタの収束速度向上をはかる。

本論文は5章より構成されている。以下に各章の概要を述べる。まず第2章では、本研究全体を通じて用いた基礎的な原理について述べる。また、多チャンネル音場再生システムにおける従来の逆フィルタ処理を紹介する。

第3章では、2チャンネル-2点制御系の新たな音場再生システムを提案し、実環境データを用いた計算機シミュレーションにより収束特性について検証する。

第4章では、聴覚特性に基づく重み付け出力誤差を用いて、適応フィルタの計算効率の向上をはかる。第3章で扱う実環境データを用いて、計算機シミュレーションを行うことにより、収束速度の検証を行う。

最後に、以上で得られた結果を元に第5章において本論文の結論を述べる。

第 2 章

音場再生システムと逆特性の近似

2.1 はじめに

音場再生システムについては，音の再現に必要なハードウェア構成や再生手法の違いから様々な手法が提案されてきた．本章ではまず，音場などの音響伝達系を含んだ音響システムを線形モデルとして表現されることを説明する．線形モデルは音響システムの動作の理解や解析に有用となる．次に，ラウドスピーカによる音場再生の考え方とその達成方法について述べる．本論文では，2つのラウドスピーカを用いて空間上の2つの制御点を制御する再現システムを念頭においている．そこで，多点入力-多点出力の音場再生システムにおける逆フィルタの従来の設計法とその問題点について述べる．

2.2 音響システムと線形モデル

空間内の2点（音源と受音点など）の音の伝わり方を表す関数を空間伝達関数 $G(z)$ という．空間が屋外などの自由空間である場合，空間伝達関数 $G(z)$ は音源からの距離のみに依存する．これは音圧が音源からの距離に反比例して減少する距離減衰特性と，その距離を音が伝搬するのに要する伝搬遅延時間を表すものとなっている．一方，音場が図 2.1 に示すように反射音を有する室内などの閉空間において，音源であるラウドスピーカと受音器となるマイクロホンを設置し，この2点間の音の伝達について考える．ここでラウドスピーカに入力した音を $x(t)$ ，マイクロホンで受音した信号を $y(t)$ とする．このときマイクロホンで観測される音 $y(t)$ には，音源から受音点に直接到達する音（直接音）と，壁面で反射された多

2.2 音響システムと線形モデル

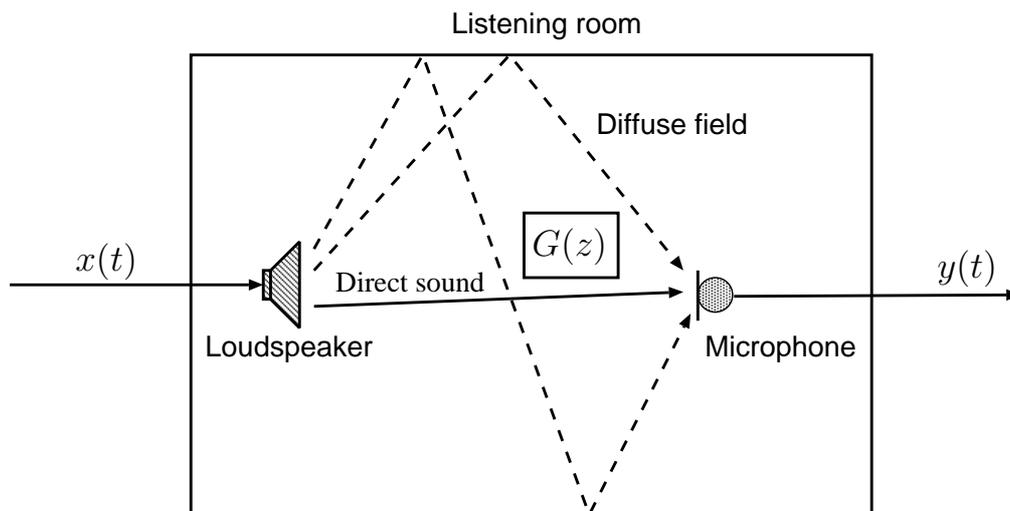


図 2.1 反射音を有する室内における音の伝搬の様子

数の反射音が含まれる．反射音はマイクロホンで受信されるまでに壁面で反射された回数によって一次反射音，2次反射音などに分類される．音の反射は壁面の吸音などによって音のエネルギーが消滅してしまうまで多数回繰り返される．音源からパルス音を発生させたときにマイクロホンで受信される信号（インパルス応答）は時間の経過と共に，まず直接音が受信され，次に1次，2次などの低次反射音が散発的に受信される．その後，続けて高次反射音が受信される．ここで低次の反射音は単に初期反射音と呼ばれ，また高次の反射音はインパルス応答長が時間の経過と共に級数的に増加し，残響音と呼ばれる場合がある．このインパルス応答を標準化して得られた離散数列を $\{g(0), g(1), g(2), \dots\}$ と表す．室内伝達関数 $G(z)$ は，この離散数列を z 変換した次式により求められる．

$$G(z) = \sum_{i=0}^{\infty} g(i)z^{-i} \quad (2.1)$$

音源，マイクロホン，空間音響伝達系による影響はすべて線形であると仮定すると， $x(t)$ と $y(t)$ の関係はインパルス応答 $g(k)$ を用いて以下のように表される．

$$y(t) = \sum_{k=0}^{\infty} g(k)x(t-k) \quad (2.2)$$

この式は，観測したインパルス応答 $g(k)$ を用いることにより，任意の入力信号 $x(n)$ に対する応答 $y(t)$ が求められることを意味している．式 (2.2) の両辺を z 変換すると，次式のよう

2.3 逆特性の近似

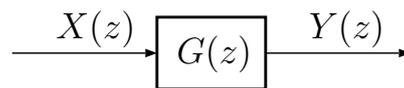


図 2.2 図 2.1 の等価回路

に変形される．

$$Y(z) = G(z)X(z) \quad (2.3)$$

ただし， $X(z)$ ， $Y(z)$ はそれぞれ $x(t)$ ， $y(t)$ の z 変換を意味する．式 (2.3) の関係より，図 2.1 の音響伝達系は図 2.2 のような等価回路で表される．このように室内音場における音響の入出力システムは，線形系としてモデル化することができる．

2.3 逆特性の近似

トランスオーラル系の音場再生システムでは，ラウドスピーカに入力される所望信号と受音点での観測信号が等しくなるように制御することを目的とする．つまり図 2.2 に示す線形モデルを想定した場合， $X(z) = Y(z)$ となるように制御しなければならない．しかし実際の音場においては，受音点で観測される音は式 (2.3) に示したように伝達特性 $G(z)$ の影響を受けるため，この影響を補正する必要がある．このような影響を $D(z)$ から除去し， $X(z)$ を復元するために，図 2.3 に示すように補正フィルタ $C(z)$ を用いる．ここで補正フィルタとは空間伝達特性の逆特性に近似したフィルタである．今，この補正フィルタ $C(z)$ が伝達特性 $G(z)$ の完全な逆特性を示すものであると仮定すると，

$$C(z) = G^{-1}(z) \quad (2.4)$$

なる関係が成り立つ．図 2.3 の関係から補正フィルタと伝達特性を畳み込むことによって

$$\begin{aligned} G(z)C(z) &= G(z)G^{-1}(z) \\ &= 1 \end{aligned} \quad (2.5)$$

となる．したがって伝達特性が 1 を示すとき，受音点では所望信号 $x(t)$ がそのまま再生されることになる．ただしこのような関係が成り立つためには伝達関数 $G(z)$ が最小位相系を

2.3 逆特性の近似

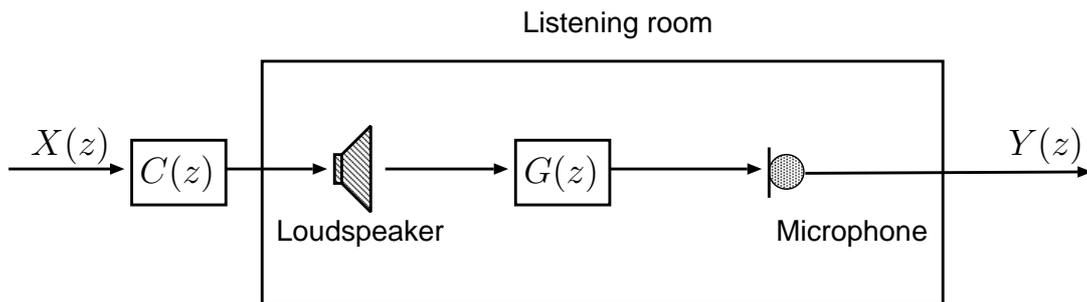


図 2.3 室内伝達特性の補正

示す ($G(z)$ に対するすべての零点が単位円内に含まれる) 必要がある。しかしながら一般の室内においてはその多くが非最小位相系であることが知られており、したがって正確な逆特性を推定することは極めて困難であるため式 (2.5) のような関係は実現できない。

非最小位相系に対する逆特性を実現するための方法として、安定性が保証されている FIR フィルタを用いて伝達特性の逆特性を近似的に求める方法が提案されている。代表的な手法として、逆特性を誤差エネルギーを最小にする FIR フィルタで近似する手法が挙げられる。これは

$$E(z) = 1 - G(z)C(z) \quad (2.6)$$

で与えられる誤差 $E(z)$ のエネルギー ($E(z)$ の 2 乗) を最小とする $C(z)$ の探索問題に帰着する。 $G(z)$ が非最小位相系である場合、そのまま $G(z)$ の逆特性 $1/G(z)$ を近似すると、因果性や安定性を満たさないフィルタとして設計されるため、フィルタを安定化させるような様々な工夫がなされている。具体的には不要となる零点を除去したり、遅延や雑音を付加することすることが提案されている。

一方、MINT 理論によると、複数の空間伝達経路を設けることにより、特性が最小位相を示さない場合でも正確な逆特性を実現できることが示されている。図 2.4 に MINT 理論に基づく音場制御系を示す。受音点における観測信号 $Y(z)$ は

$$\begin{aligned} D(z) &= G_1(z)C_1(z)X(z) + G_2(z)C_2(z)X(z) \\ &= (G_1(z)C_1(z) + G_2(z)C_2(z))X(z) \end{aligned} \quad (2.7)$$

と表される。この系で逆フィルタ処理を行うには、次の関係を満たす FIR フィルタ $C_1(z)$ 、

2.4 多チャンネル音場再生システムにおける従来の逆特性近似

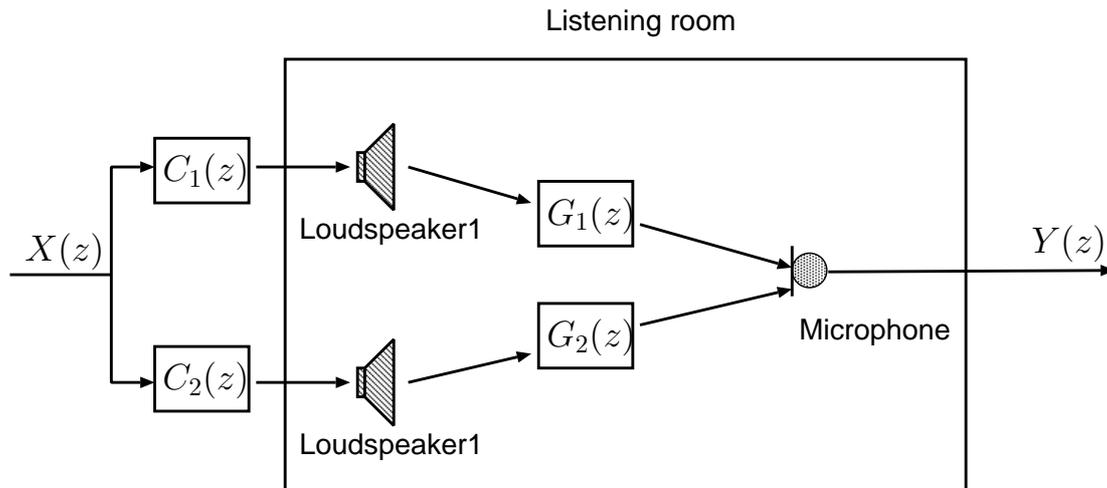


図 2.4 MINT 理論に基づく音場制御系

$C_2(z)$ を求めなければならない。

$$G_1(z)C_1(z) + G_2(z)C_2(z) = 1 \quad (2.8)$$

この解の存在条件は、 $G_1(z)$ 、 $G_2(z)$ が有限次数であり、かつ共通の零点を持たないことである。これらの条件が満たされている場合、式 (2.8) を満たす有限次数の $C_1(z)$ 、 $C_2(z)$ が存在することは、代数学の Diophantos の方程式として証明されている。

2.4 多チャンネル音場再生システムにおける従来の逆特性近似

本節ではまず MINT 理論に基づく多チャンネル-多点制御系の設計指針を述べ、具体的な設計方法を示した後に、従来法の問題点について述べる。

2.4.1 多チャンネル-多点制御近似逆フィルタの設計方針

図 2.5 に 2 次音源数 I 、制御点数 J の多チャンネル音場再生系において逆特性を近似するための近似逆フィルタ（補正フィルタ）群の構成を示す。ここで室内伝達関数のインパルス応答を $G_{ji}(z)$ 、補正フィルタを $C_{ij}(z)$ 、所望信号を $X_j(z)$ 、観測信号を $Y_j(z)$ とする。また $i(=1, 2, \dots, I)$ と $j(=1, 2, \dots, J)$ はそれぞれ 2 次音源番号、制御点番号である。多チャネ

2.4 多チャンネル音場再生システムにおける従来の逆特性近似

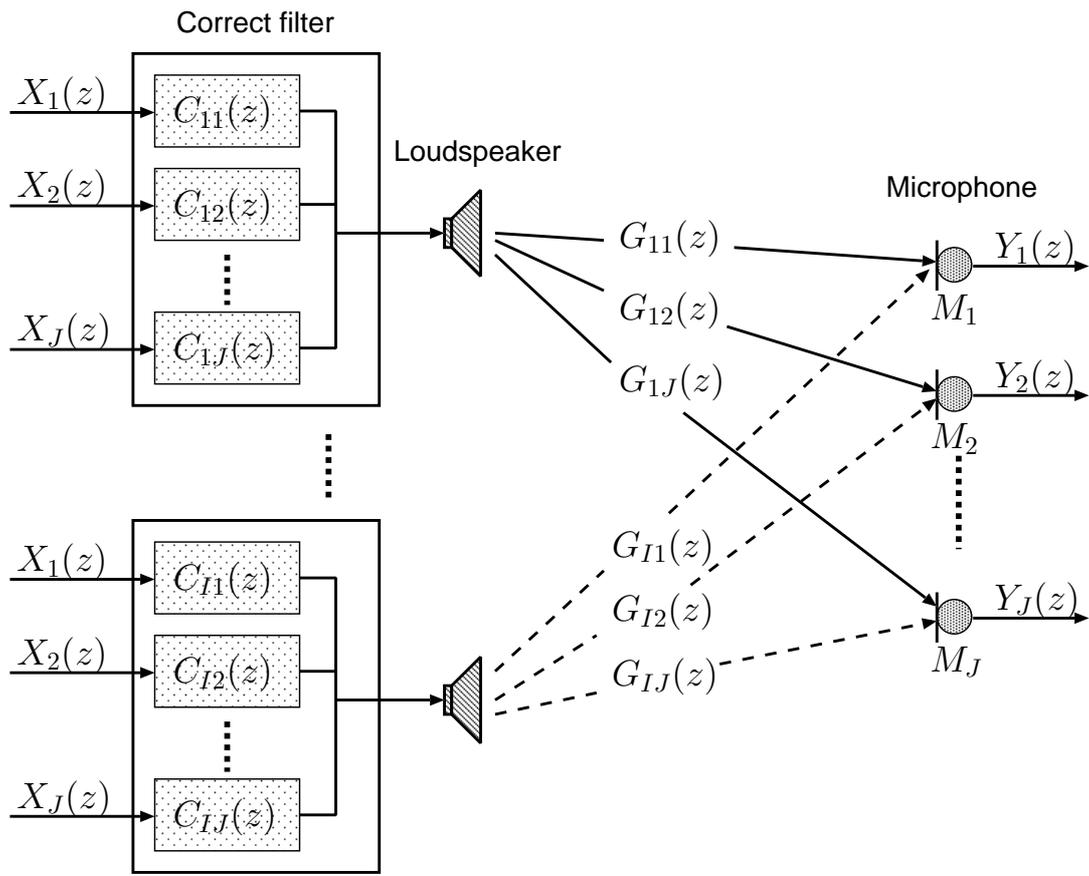


図 2.5 2 次音源数 I , 制御点数 J の音響系

ル音場再生システムでは，各制御点に対する制御を他の点とは独立に行われることが望まれる．そこで補正フィルタ群においては，入力信号 X_j をパルスとしたときに制御点 M_j へのみパルスが観測されるような補正フィルタ群を基本単位とする．すなわち

$$G(z)C(z) = I_I \quad (2.9)$$

を満たす $C(z)$ を求めればよい．ただし，

$$G(z) = \begin{bmatrix} G_{11}(z) & G_{12}(z) & \cdots & G_{1I}(z) \\ G_{21}(z) & G_{22}(z) & \cdots & G_{2I}(z) \\ \vdots & \vdots & \ddots & \vdots \\ G_{J1}(z) & G_{J2}(z) & \cdots & G_{JI}(z) \end{bmatrix} \quad (2.10)$$

2.4 多チャネル音場再生システムにおける従来の逆特性近似

$$C(z) = \begin{bmatrix} C_{11}(z) & C_{12}(z) & \cdots & C_{1I}(z) \\ C_{21}(z) & C_{22}(z) & \cdots & C_{2I}(z) \\ \vdots & \vdots & \ddots & \vdots \\ C_{I1}(z) & C_{I2}(z) & \cdots & C_{IJ}(z) \end{bmatrix} \quad (2.11)$$

とし、 I_I は $I \times I$ の大きさの単位行列とする。

各制御点に対して上記の補正フィルタ係数を求め、それらを2次音源毎に畳み込むことによって逆特性を近似する制御系が構成される。

2.4.2 MINT 理論に基づく近似逆フィルタ設計

多チャネル音場再生を実現するための近似逆フィルタ設計として、まずその基礎となる図2.3に示すような1入力-1出力の再生系を対象とする近似逆フィルタ設計法について述べる。

ここでは入力信号 $X(z)$ が単位パルス信号であるとする。また対象とする伝達系 $G(z)$ のインパルス応答を $g(k)$ と表す。 $g(k)$ の値は、時刻 L_g 以降、零とみなせるものと仮定する。つまり、 $g(k)$ は $0 \leq k \leq L_g$ でのみ非零の値をもつものとする。近似逆フィルタ(補正フィルタ) $C(z)$ は L 次の FIR フィルタを用いるものとし、そのインパルス応答(フィルタ係数)を $c(k)$ 、 $k = 0, 1, \dots, L$ と表す。このとき、対象とする伝達系と近似逆フィルタを縦続に接続して、単位サンプル信号を入力した場合に制御点で観測される信号 $y(k)$ は、 $g(k)$ と $c(k)$ の畳み込みとして次式のように表される。

$$y(k) = \sum_{i=0}^L g(k-i)c(i) \quad (2.12)$$

ここで $k < 0, k > L_g$ のとき $g(k) = 0$ となり、 $k < 0, k > L$ のとき $c(k) = 0$ であるので、 $y(k)$ は $0 \leq k \leq L_g + L$ でのみ非零の値をもつ。このことより式(2.12)の関係は以下のよ

2.4 多チャネル音場再生システムにおける従来の逆特性近似

うな行列で表すことができる．

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ \vdots \\ \vdots \\ \vdots \\ y(L_g + L) \end{bmatrix} = \begin{bmatrix} g(0) & 0 & 0 & \cdots & 0 \\ g(1) & g(0) & 0 & \cdots & 0 \\ g(2) & g(1) & g(0) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g(L_g) & g(L_g - 1) & g(L_g - 2) & \cdots & 0 \\ 0 & g(L_g) & g(L_g - 1) & \cdots & g(0) \\ 0 & 0 & g(L_g) & \cdots & g(1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & g(L_g) \end{bmatrix} \begin{bmatrix} c(0) \\ c(1) \\ c(2) \\ \vdots \\ \vdots \\ \vdots \\ c(L) \end{bmatrix} \quad (2.13)$$

$y(k)$ を要素とした左辺のベクトルを \mathbf{y} , $g(k)$ を要素とした右辺の行列を \mathbf{g} , $c(k)$ を要素とした右辺のベクトルを \mathbf{c} とすると以下のように書き換えられる．

$$\mathbf{y} = \mathbf{g}\mathbf{c} \quad (2.14)$$

$y(k)$ は $x(k)$ と同値で観測されることが望まれる．したがって計算条件より \mathbf{y} は単位パルス信号であり , $\mathbf{y} = [1, 0, \dots, 0]^T$ となる．行列 \mathbf{g} が正方行列であり , かつ正則である場合 , 以下のように \mathbf{g} の逆行列を求めることによって \mathbf{c} が計算される．

$$\mathbf{c} = \mathbf{g}^{-1}\mathbf{y} \quad (2.15)$$

\mathbf{g} が正則でない場合 , 以下のような最小 2 乗解を用いることにより \mathbf{c} を求めることができる .

$$\mathbf{c} = (\mathbf{g}^T \mathbf{g})^{-1} \mathbf{g}^T \mathbf{y} \quad (2.16)$$

ただし \mathbf{g}^T は \mathbf{g} の転置行列である . これは式 (2.6) の考え方と一致する . 上記の考えに基づけば , 多チャネル-多点制御系における近似逆フィルタ設計を以下のような行列で定式化で

2.5 従来法の問題点

きる .

$$\begin{bmatrix} y_1(0) \\ y_1(1) \\ \vdots \\ y_1(L_g) \\ \hline \vdots \\ y_J(0) \\ y_J(1) \\ \vdots \\ y_J(L_g) \end{bmatrix} = \begin{bmatrix} g_{11}(0) & \cdots & 0 & \cdots & g_{I1}(0) & \cdots & 0 \\ g_{11}(1) & \cdots & 0 & \cdots & g_{I1}(1) & \cdots & 0 \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & g_{11}(L_g) & \cdots & 0 & \cdots & g_{I1}(L_g) \\ \hline \vdots & & & & \vdots & & \\ \hline g_{1J}(0) & \cdots & 0 & \cdots & g_{IJ}(0) & \cdots & 0 \\ g_{1J}(1) & \cdots & 0 & \cdots & g_{IJ}(1) & \cdots & 0 \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & g_{1J}(L_g) & \cdots & 0 & \cdots & g_{IJ}(L_g) \end{bmatrix} \begin{bmatrix} c_1(0) \\ c_1(1) \\ \vdots \\ c_1(L) \\ \hline \vdots \\ c_J(0) \\ c_J(1) \\ \vdots \\ c_J(L) \end{bmatrix} \quad (2.17)$$

MINT 理論によれば $I > J$ かつ g が正方行列となる条件，すなわち以下の式を満たす場合，近似逆フィルタが厳密な FIR フィルタとして設計できることが保証されている .

$$J(L + L_g - 1) = IL \quad (2.18)$$

2.5 従来法の問題点

ここでは，多点制御音場再生における従来の近似逆フィルタ設計手法についてその問題点を述べる .

MINT 理論に基づいた設計手法によって，近似逆フィルタは厳密な FIR フィルタとして求めることは可能である . しかし通常の部屋等を再生空間とした場合その残響時間は長く，すなわち空間伝達関数のインパルス応答長は長くなってしまふ . また伝達特性の影響を制御するための系の規模を大きくすると行列 g もそれに比例して大きくなってしまふ . したがって計算機が要求する計算量とメモリ量は膨大なものとなり，計算機での実現は困難なものとなる . このようなことから，従来の MINT 理論による近似逆フィルタを算出することは理論的に可能ではあるが，現実的ではないといえる .

2.6 まとめ

本章では，本論文で取り扱う音場再生システム，および従来の近似逆フィルタ設計手法についてまとめた．MINT 理論に従った近似逆フィルタは厳密な FIR フィルタを設計しうる反面，計算量の多さから実用が困難であることを指摘した．そこで比較的少ない計算量で近似逆フィルタを算出する制御系を確立することが望まれる．この問題に対する提案を第 3 章で行う．

第 3 章

2 チャンネル-2 点制御による多入力信号補正法

3.1 はじめに

ラウドスピーカ再生による多チャンネル音場再生システムにおいては室内伝達特性の影響を除去する近似逆フィルタの設計が不可欠であり，そのための方法として MINT 理論に基づく設計手法が提案されている [2][3]．しかしながら，前章で述べたとおり空間伝達関数のインパルス応答長が長い場合，計算機が要求する計算量とメモリ量は膨大となり，実用化に至っていない．そこで本報告では制御点数 2 に対し 2 次音源数 2 で構成される新たな制御系を提案する．さらに実環境で計測した空間伝達関数を用いて計算機シミュレーションを行い収束特性について評価する．

3.2 2 チャンネル-2 点制御系における問題点

ここでは 2 個の 2 次音源を用いて 2 個の制御点を制御する音響系に対して線形モデル化を行い，2 次音源に入力される原音と制御点で観測される観測信号が一意に等化できないことについて述べる．

図 3.1 に 2 チャンネル-2 点制御系モデルを示す．対象とする伝達系 $G_{ij}(z)$ のインパルス応答を $g_{ij}(k)$ と表す． g_{ij} の値はいずれも，時刻 L_g 以降，零とみなせるものと仮定するつまり， $g(k)$ は $0 \leq k \leq L_g$ でのみ非零の値をもつものとする．また，近似逆フィルタ（補正

3.2 2チャンネル-2点制御系における問題点

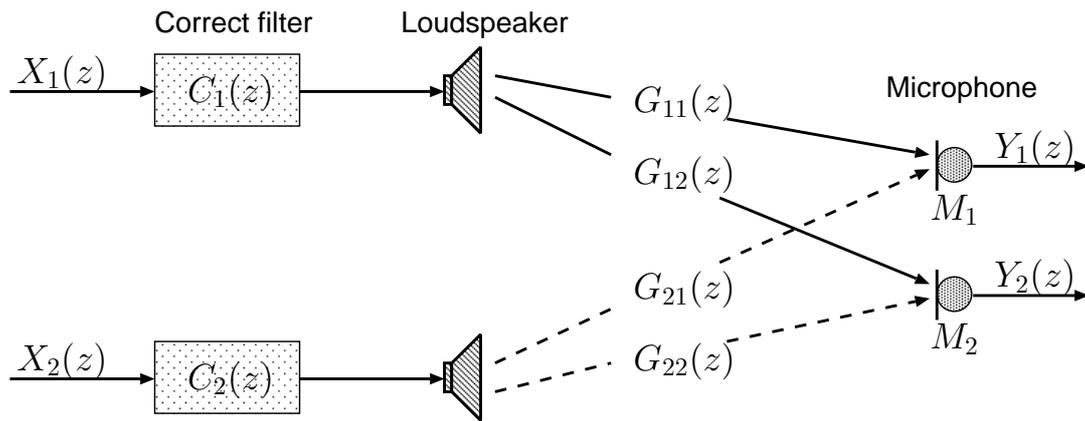


図 3.1 2チャンネル-2点制御系線形モデル

フィルタ) $C_i(z)$ は L 次の FIR フィルタでそのインパルス応答を $c(k)$, $k = 0, 1, \dots, L$ と表す. 入力される原音は $x_i(t)$ と表し, 観測信号は $y_j(t)$ と表す. ここで補正フィルタ $C_i(z)$ が 1 を示す (補正しない) 場合を考える. このとき, 2 つの制御点 M_j で観測される信号 $y_j(t)$ と原音 $x_i(t)$ の関係はインパルス応答 $g_{ij}(k)$ を用いて以下のように表される.

$$y_1(t) = \sum_{k=0}^{L_g} g_{11}(k)x_1(t-k) + \sum_{k=0}^{L_g} g_{21}(k)x_2(t-k) \quad (3.1)$$

$$y_2(t) = \sum_{k=0}^{L_g} g_{22}(k)x_2(t-k) + \sum_{k=0}^{L_g} g_{12}(k)x_1(t-k) \quad (3.2)$$

ここで式 (3.1) における右辺の第 1 項は, 右側 2 次音源から右側制御点に至る成分を示しており, 第 2 項は本来, 左側 2 次音源から左側観測信号に伝えられるべき音が右側観測信号へ伝達されるクロストーク現象を示している. したがって, 伝達特性 $G_{11}(z)$ に対する近似逆フィルタを算出し, これを補正フィルタ $C_1(z)$ として使用しても, 右チャンネルから G_{21} を通過するクロストーク信号が加算されることにより右側観測信号は右側原信号とは大きく異なってしまふ. また, 左右のチャンネル対して制御系は対称な構成であるため, 式 (3.2) についても例外ではない.

以上のことから 1 つの補正フィルタは以下に示す 2 つの要求を同時に実現するような機能を持たなければならない.

3.3 多入力信号補正システム

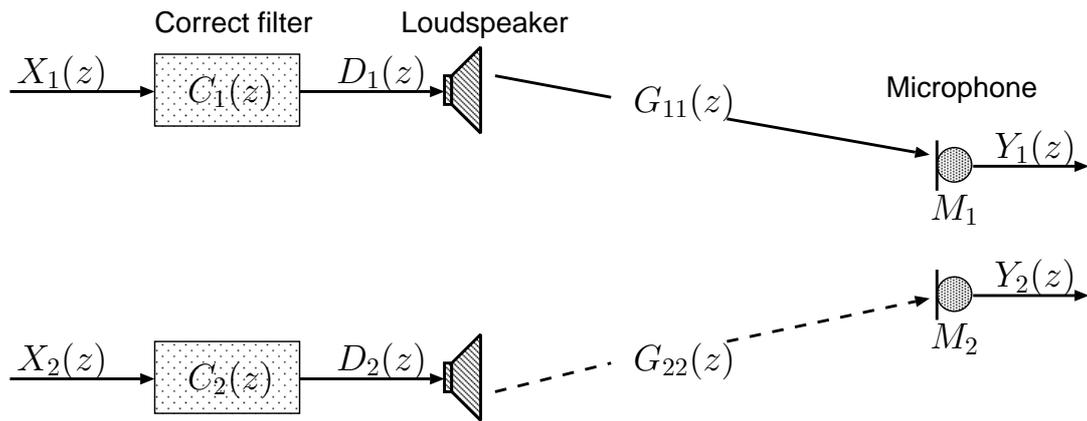


図 3.2 2 チャンネル-2 点制御系線形モデル (クロストーク無し)

- 伝達特性の影響を軽減させること
- クロストークの影響を軽減させること

3.3 多入力信号補正システム

本節ではまず，伝達特性の影響を軽減させるために，適応フィルタを用いて近似的に逆特性を推定する仮定について述べ，クロストーク現象による影響を軽減させることへの対策について述べる．また，本論文で提案する 2 チャンネル-2 点制御系を今後は多入力信号補正システムと呼ぶ．

3.3.1 逆特性の近似的推定法

本節では，伝達特性の影響を軽減させる方法として適応フィルタを用いて，逆特性を近似的に導出する仮定について述べる．説明簡単化のため，ここでは図 3.2 に示すような，クロストークの影響がない線形モデルを考える．

原音である所望信号 $x_i(t)$ が補正フィルタ $C_i(z)$ 通過後の信号 $d_i(t)$ は次のように表される．

$$d_i(t) = \mathbf{c}_{i,L} \mathbf{x}_{i,L}^T(t) \quad (3.3)$$

3.3 多入力信号補正システム

ここで

$$\mathbf{c}_{i,L} = [c_{i,0}, c_{i,1}, \dots, c_{i,L}] \quad (3.4)$$

$$\mathbf{x}_{i,L}(t) = [x_i(t), x_i(t-1), \dots, x_i(t-L)] \quad (3.5)$$

となる．また， L は補正フィルタのインパルス応答長である．このとき観測信号 $y_j(t)$ はクロストークがないものと仮定すると

$$y_j(t) = \mathbf{d}_{i,L_g}(t) \mathbf{g}_{ij,L_g}^T \quad (3.6)$$

となる．ただし

$$\mathbf{d}_{i,L_g}(t) = [d_i(t), d_i(t-1), \dots, d_i(t-L_g)] \quad (3.7)$$

$$\mathbf{g}_{ij,L_g} = [g_{ij}(0), g_{ij}(1), \dots, g_{ij}(L_g)] \quad (3.8)$$

である．ここで，補正フィルタ $C_i(z)$ が伝達関数 $G_{ij}(z)$ に対する近似逆フィルタと同様の性質であると仮定すると， z 変換による表現で以下の近似式が成り立つ．

$$\begin{aligned} Y_j(z) &= G_{ij}(z) D_i(z) \\ &= G_{ij}(z) C_i(z) X_i(z) \\ &\approx G_{ij}(z) G_{ij}^{-1}(z) X_i(z) \\ &= X_i(z) \end{aligned} \quad (3.9)$$

これは，伝達関数に対する近似逆フィルタを用いることで原音である所望信号と制御点での観測信号を等化できる可能性があることを示している．図 3.3 に近似逆フィルタ係数を推定するための構成図を示す．伝達特性 g_{ij,L_g} の逆特性を推定するための適応フィルタ係数を $\mathbf{h}_{ij,L}(t)$ ，適応フィルタの入力状態ベクトルを $\mathbf{y}_{ij,L}(t)$ とすると，入出力関係は

$$s_{ij}(t) = \mathbf{h}_{ij,L}^T(t) \mathbf{y}_{ij,L}(t) \quad (3.10)$$

で与えられる．ただし

$$\mathbf{h}_{ij,L}(t) = [h_{ij,0}(t), h_{ij,1}(t), \dots, h_{ij,L}(t)] \quad (3.11)$$

$$\mathbf{y}_{ij,L}(t) = [y_j(t), y_j(t-1), \dots, y_j(t-L)] \quad (3.12)$$

3.3 多入力信号補正システム

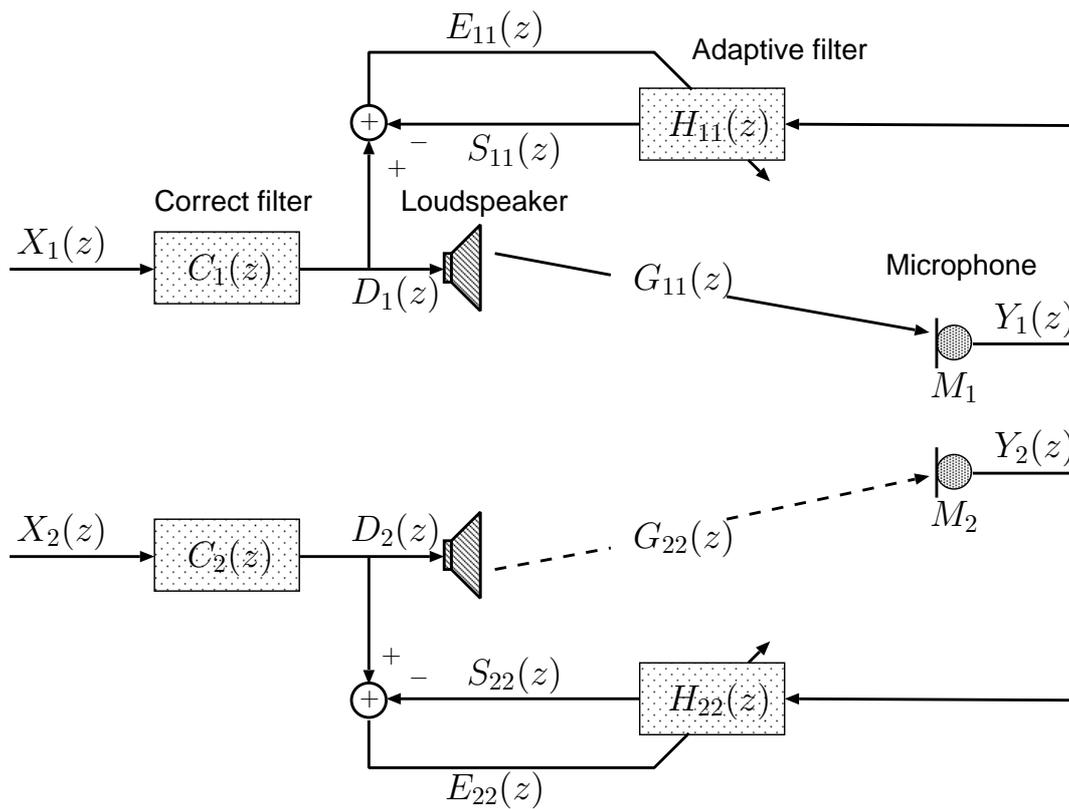


図 3.3 近似逆フィルタと構成図

である．一方，系への入力信号を式 (3.3) で与えられるものとする，出力誤差 $e_{ij}(t)$ は

$$e_{ij}(t) = d_i(t) - s_{ij}(t) \quad (3.13)$$

となる．ここで，出力誤差が $e(t) = 0$ の場合について考えると

$$\mathbf{y}_{j,L}(t) \mathbf{h}_{ij,L}^T(t) = \mathbf{x}_{i,L}(t) \mathbf{c}_{j,L}^T(t) \quad (3.14)$$

となる．室内における伝達特性の近似逆フィルタとなる適応フィルタ係数 $\mathbf{h}_{ij,L}$ を補正フィルタ係数 $\mathbf{c}_{ij,L}$ として与えると，式 (3.14) は

$$\begin{aligned} \mathbf{y}_{j,L}(t) \mathbf{h}_{ij,L}^T(t) &= \mathbf{x}_{i,L}(t) \mathbf{h}_{ij,L}^T(t) \\ \mathbf{y}_{j,L}(t) &= \mathbf{x}_{i,L}(t) \end{aligned} \quad (3.15)$$

となる．これより，入力誤差 $e(t) = 0$ であるならば，補正フィルタへの入力信号 $x_j(t)$ と観測信号 $y_j(t)$ は等しくなる．したがって，入力誤差 $e_{ij}(t)$ が最小となるように適応フィルタ

3.3 多入力信号補正システム

の係数を更新していく必要がある．そこで係数更新には

$$\mathbf{h}_{ij,L}(t+1) = \mathbf{h}_{ij,L}(t) + \alpha \frac{\mathbf{y}_{j,L}(t)}{\|\mathbf{y}_{j,L}(t)\|^2} e_{ij}(t) \quad (3.16)$$

で示される学習同定法を用いる [18]．ここで， $\mathbf{y}_{j,L}(t)$ は入力信号の状態ベクトル， $\mathbf{h}_{ij,L}(t)$ は適応フィルタのフィルタ係数であり， $\|\cdot\|$ はベクトルのユークリッドノルムを表す．また， α はステップゲインと呼ばれ，フィルタ係数修正の大きさを制御するパラメータである．

3.3.2 クロストークの影響を考慮した補正フィルタ係数の導出

本節では 3.3.1 の伝達特性の逆特性推定法を利用して，クロストークの影響を軽減させる多入力信号補正システムについて提案する．2 つの 2 次音源から受聴者の左右耳を模した 2 つ制御点に至るまでの伝達経路は図 3.1 より 4 本存在する．そのため左右の制御点において受音される観測信号 $y_1(t)$ ， $y_2(t)$ はそれぞれ以下に示す式で与えられる．

$$y_1(t) = \mathbf{x}_{1,L_g}(t) \mathbf{g}_{11,L_g}^T + \mathbf{x}_{2,L_g}(t) \mathbf{g}_{21,L_g}^T \quad (3.17)$$

$$y_2(t) = \mathbf{x}_{2,L_g}(t) \mathbf{g}_{22,L_g}^T + \mathbf{x}_{1,L_g}(t) \mathbf{g}_{12,L_g}^T \quad (3.18)$$

ここで式 (3.17)，(3.18) の第 2 項はクロストーク成分を表している．すなわち，2 チャネルの音場再生システムではこのクロストークの影響により式 (3.15) を成立させることが難しくなっている．この解決策として，2 本のクロストークにおける伝達関数 g_{12,L_g} と g_{21,L_g} は互いに強い相関を持つことに注目し，補正フィルタ係数はクロストークとそれ以外のパスの伝達特性の逆特性を用いて導出することを考える．図 2 に 2 チャネル-2 点制御による音場再生を実現するための構成ブロック図を示す．

ここでまず，右側入力信号 $d_1(t)$ と適応フィルタ $H_{11}(z)$ を通過する信号 $s_{11}(t)$ から出力誤差 e_{11} を算出している．つまり時間を追ってこの処理が繰り返され 3.3.1 で述べた原理で，出力誤差 e_{11} は次第に小さくなる．したがって，適応フィルタ $H_{11}(z)$ は右側 2 次音源から右側制御点に至る伝達特性 g_{11,L_g} の逆特性を近似するフィルタとなる． $H_{11}(z)$ で得られたフィルタ係数を右側補正フィルタ $C_1(z)$ に与えることによって，右側観測信号 $y_1(t)$ は右側所望信号 $x_1(t)$ に近似させることができる．しかし，この補正処理がクロストークによ

3.3 多入力信号補正システム

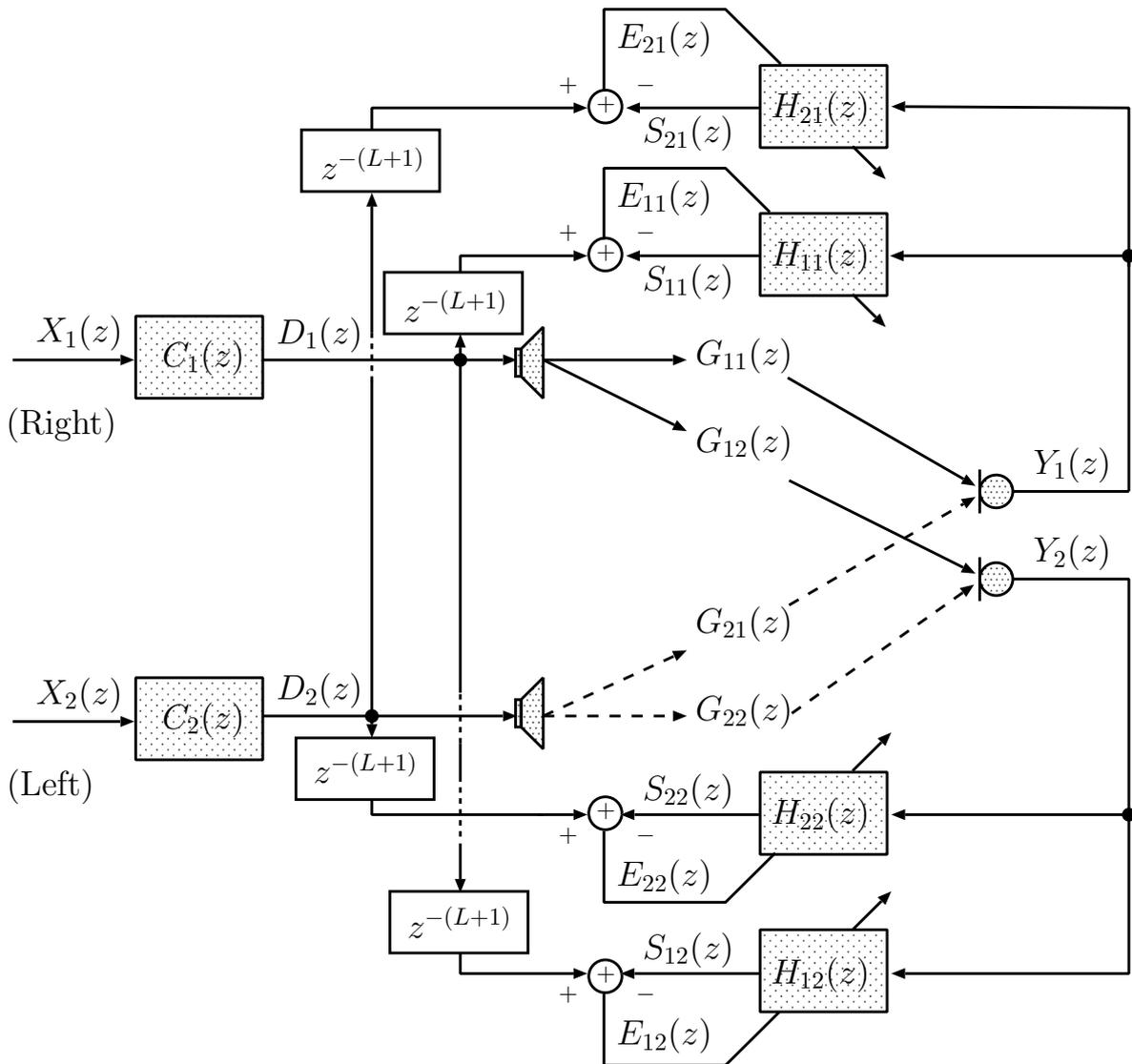


図 3.4 2 チャンネル音場再生システムの構成図

り左側観測信号 $y_2(t)$ を右側所望信号 $x_1(t)$ に近似させるように作用するため，これを防ぐ工夫が必要となる．

ところで，適応フィルタ $H_{21}(z)$ は左側入力信号 $d_2(t)$ とフィルタの出力信号 $s_{21}(t)$ から出力誤差を算出しているため，左側 2 次音源から右側制御点に至るクロストーク成分の特性 g_{21,L_g} の逆特性を近似するフィルタとなる．通常， $H_{21}(z)$ 得られるフィルタ係数は，右側補正フィルタ $C_1(z)$ と接続関係にないため，フィルタ係数 c_1 として与えることはできない．しかし，左右の制御点間の距離はせいぜい人間の頭部における左右の耳介距離であるため，

3.3 多入力信号補正システム

クロストークにおける伝達特性 g_{12,L_g} と g_{21,L_g} は互いに強い相関をもつと考えられる．このことを利用すると，適応フィルタ $H_{21}(z)$ は伝達関数 g_{12,L_g} の近似逆特性を推定する予測フィルタとして代用することができる．すなわち，この予測フィルタ係数を右側補正フィルタ係数 c_1 として与えることで， g_{12,L_g} を通過するクロストーク成分を左側所望信号 $x_2(t)$ に近似させることができる．

したがって， $H_{11}(z)$ で得られたフィルタ係数と $H_{21}(z)$ で得られたフィルタ係数を合成させたものを右側補正フィルタ係数として与えることによって，それぞれの伝達特性の影響を除去すると同時に，クロストークの影響を軽減させる働きをもつ補正フィルタが得られる．ここまでは右側チャンネルに対する対策であったが，左側チャンネルについても同じ構成であるため g_{21,L_g} を通過するクロストーク成分を右側所望信号 $x_1(t)$ に近似させることができる．ここで左右の補正フィルタ係数はそれぞれ以下の式で与えられる．

$$c_{1,L}(t+1) = w_1 h_{11,L}(t) + (1 - w_1) h_{21,L}(t) \quad (3.19)$$

$$c_{2,L}(t+1) = w_2 h_{22,L}(t) + (1 - w_2) h_{12,L}(t) \quad (3.20)$$

ただし， w_i は第 1 項と第 2 項に対する重み付けパラメータ ($0 < w_i \leq 1$) である．

ここで提案してきた 2 チャンネル-2 点制御系による多入力信号補正法の収束性については数学的に証明されていないが，次節以降に示すとおり，収束性が実験的に確認されている．

3.3.3 補正フィルタ係数の更新手順

本節では補正フィルタ係数を導出するために，学習同定法に基づいた更新手順について述べる．以下の手順手順 1- 3 を信号 1 サンプル毎に繰り返し処理することによって，右側 2 次音源から伝達特性 $G_{12}(z)$ を通過するクロストーク信号が左側所望信号 $x_2(t)$ に近似していくと考えられる．またもう一方 $G_{21}(z)$ を通過するクロストーク信号についても同様のことが言える．

手順 1: 4 つの近似逆フィルタに対する出力誤差 $e_{ij}(t)$ をそれぞれ算出する．

$$e_{ij}(t) = d_i(t) - s_{ij}(t)$$

3.4 計算機シミュレーションによる評価

手順 2: 各逆フィルタの係数ベクトルを学習同定法によってそれぞれ算出する .

$$\mathbf{h}_{ij,L}(t+1) = \mathbf{h}_{ij,L}(t) + \alpha \frac{\mathbf{y}_{ij,L}(t)}{\|\mathbf{y}_{ij,L}(t)\|^2} e_{ij}(t)$$

ただし, α はステップゲインである .

手順 3: 左右の所望信号に対する補正フィルタ係数ベクトルを次のように更新する .

$$\mathbf{c}_{1,L}(t+1) = w_1 \mathbf{h}_{11,L}(t) + (1 - w_1) \mathbf{h}_{21,L}(t)$$

$$\mathbf{c}_{2,L}(t+1) = w_2 \mathbf{h}_{22,L}(t) + (1 - w_2) \mathbf{h}_{12,L}(t)$$

ただし, w_i は重み付けパラメータ ($0 < w_i \leq 1$) である .

3.4 計算機シミュレーションによる評価

本節では, まず, 2 個の制御点を 2 個の 2 次音源を用いて制御する音場再生システムを想定し, 実環境で測定した空間伝達関数を用いて計算機シミュレーションを行った . 次に再現精度の改善量を算出し, 収束特性を比較した .

3.4.1 室内伝達関数の測定

高知工科大学のセミナールーム A511 において以下のようにインパルス応答を測定した . ラウドスピーカ, マイクロホンの配置を図 3.5~ 3.7 に示すように変更したものを 3 パターン用意した . インパルス応答の測定条件を表 3.1 に示す . 測定方法としてラウドスピーカに入力する白色雑音とマイクロホンで収録した観測信号から最小 2 乗法を用いてインパルス応答を測定した . 環境 1 について測定したインパルス応答を図 3.8~ 3.11 に示す .

3.4.2 シミュレーション条件

原音である所望信号 $x_i(t)$ として成人男性の声を 2 個のマイクロホンでステレオ録音し, 8kHz でサンプリングした音声信号 (再生時間 37.5 秒) を与える . 図 3.12, 3.13 にシミュレーションで用いた原信号を示す . また, 適応フィルタおよび補正フィルタのインパルス応

3.4 計算機シミュレーションによる評価

表 3.1 インパルス応答の測定条件

インパルス応答長	512 点
入力信号	白色雑音
サンプリング周波数	8000 Hz
量子化ビット数	16 bit

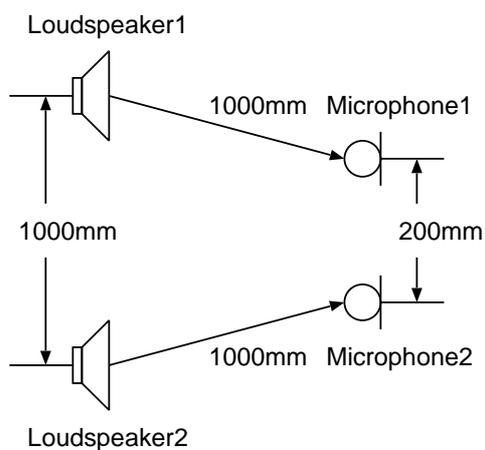


図 3.5 室内インパルス応答の測定環境 1

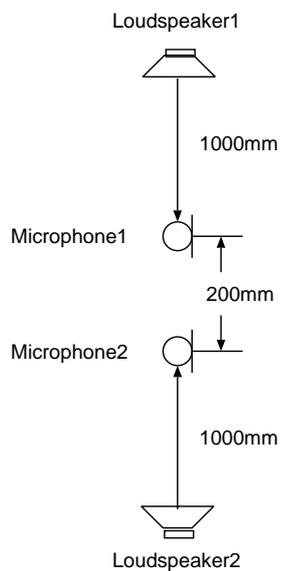


図 3.6 室内インパルス応答の測定環境 2

3.4 計算機シミュレーションによる評価

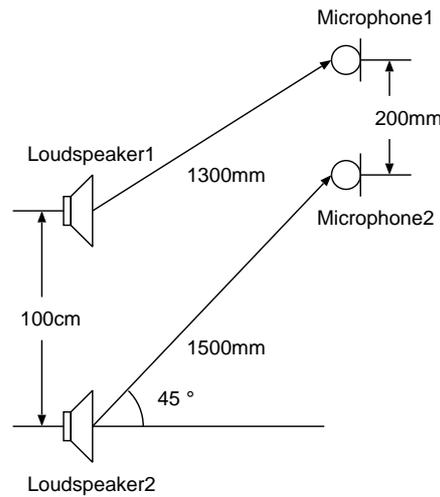


図 3.7 室内インパルス応答の測定環境 3

答長は $L = 512$, 想定する再生空間の伝達関数は 3.4.1 で測定したインパルス応答を用い , 応答長は $L_g = 512$ とした .

原音に対する再現音の再現精度は , 次のような SNR の形で表現する .

$$SNR \text{ [dB]} = 10 \log_{10} \frac{\sum [x_i(t)]^2}{\sum [y_j(t) - x_i(t)]^2} \quad (3.21)$$

ここで $x_i(t)$ は所望信号の音圧 , $y_j(t)$ は制御点 M_j における観測信号の音圧を表す .

ただし , 本論文では左右の所望信号に対して一切補正を行わずに算出した SNR と提案手法により算出した SNR との差を改善量として評価を行う . 図 3.14 はあるパラメータを設定したときの SNR と改善量の収束特性である .

3.4.3 シミュレーション結果

環境 1 のモデルで測定した室内伝達関数を与え , 学習同定法で用いるステップゲイン α , 重み付けパラメータ w_i をそれぞれ $\alpha = 0.02, 0.03, 0.04, 0.05$, $0 < w_i \leq 1$ の範囲で変動させたとき , それぞれ得られた改善量をサンプル数で平均した結果を図 3.15 , 3.16 に示す . これらの結果から , 左右のチャンネルとも良い場合で 6.5~ 7.0[dB] 程度の改善量が得られていることがわかる . しかし , 一方のチャンネルについて良好な改善量を示すようなパラメータを設定した場合 , 他方のチャンネルは収束特性の違いから良好な結果を得られない場合がある . そ

3.5 まとめ

ここで本論文では，図 3.15 と図 3.16 の改善量を各パラメータ毎に平均したもの（図 3.17 参照）について評価を行う．環境 2, 3 に対する収束特性については図 3.18, 3.19 に示す．ここで，表 3.2~ 3.4 に環境 1~ 3 においてステップゲイン $\alpha = 0.2, 0.3, \dots, 0.7$ に設定した際の最大改善量とそのときの重み付けパラメータ w_i を示す．また，以下の式で定義されるクロストーク成分の相互相関係数 r_{L_g} も示す．

$$r_{L_g} = \frac{E[g_{12,L_g}] \cdot E[g_{21,L_g}]}{\sqrt{E[g_{12,L_g}^2]} \cdot \sqrt{E[g_{21,L_g}^2]}} \quad (3.22)$$

ただし， $E[\cdot]$ は期待値を表す．

表 3.2~ 3.4 の結果より，重み付けパラメータに大きなばらつきが見られるものの，いずれの環境においてもステップゲイン $\alpha = 0.05$ 付近で最も良好な改善量が得られていることがわかる．また 2 本のクロストーク成分の相互相関係数 r_{L_g} が大きい値を示すほど良好な改善量を示していることも確認できる．これは，クロストーク成分の相互相関が高い場合，補正対象となるチャンネルの所望信号に対し，もう一方のチャンネルの所望信号に近似させるようなフィルタ係数を畳み込むことが可能となり，その結果クロストークの影響が低減されているためであると考えられる．

3.5 まとめ

本章では，クロストーク成分の相互相関に着目した多入力信号補正システムを提案した．3 種類の 2 チャンネル-2 点制御音場再生システムを想定し，実測した伝達関数を用いて計算機シミュレーションを行ったところ，いずれの再生モデルにおいても改善量が正值で収束しており，本手法が有効であることを示した．さらに，クロストーク成分の相互相関係数が大きい程，良好な結果を示すことも確認できた．

3.5 まとめ

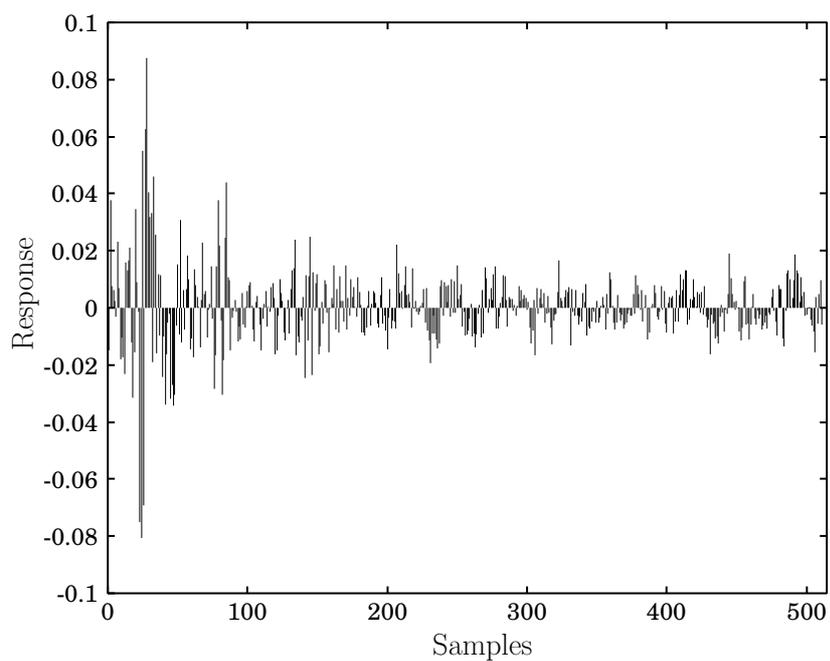


図 3.8 環境 1 の室内インパルス応答 g_{11,L_g}

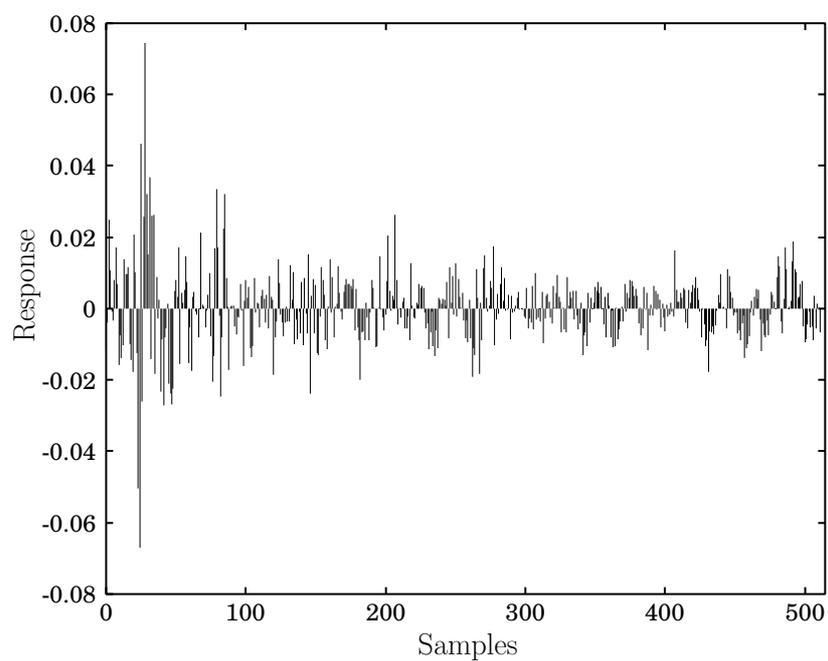


図 3.9 環境 1 の室内インパルス応答 g_{12,L_g}

3.5 まとめ

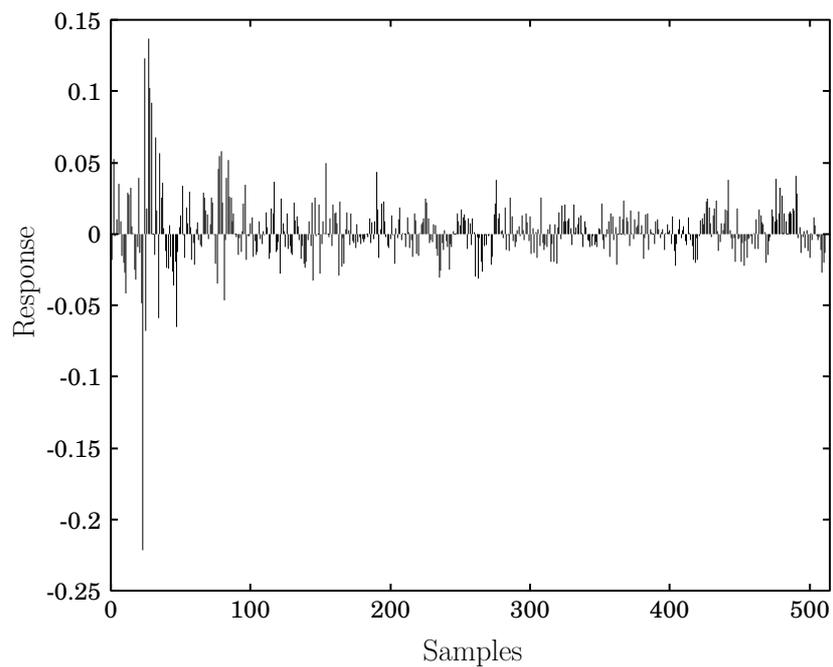


図 3.10 環境 1 の室内インパルス応答 g_{21, L_g}

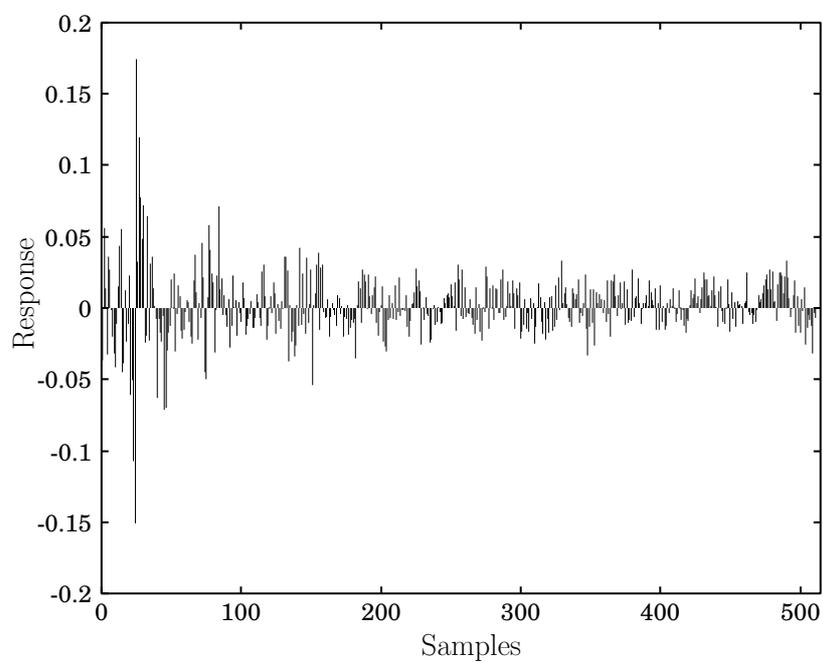


図 3.11 環境 1 の室内インパルス応答 g_{22, L_g}

3.5 まとめ

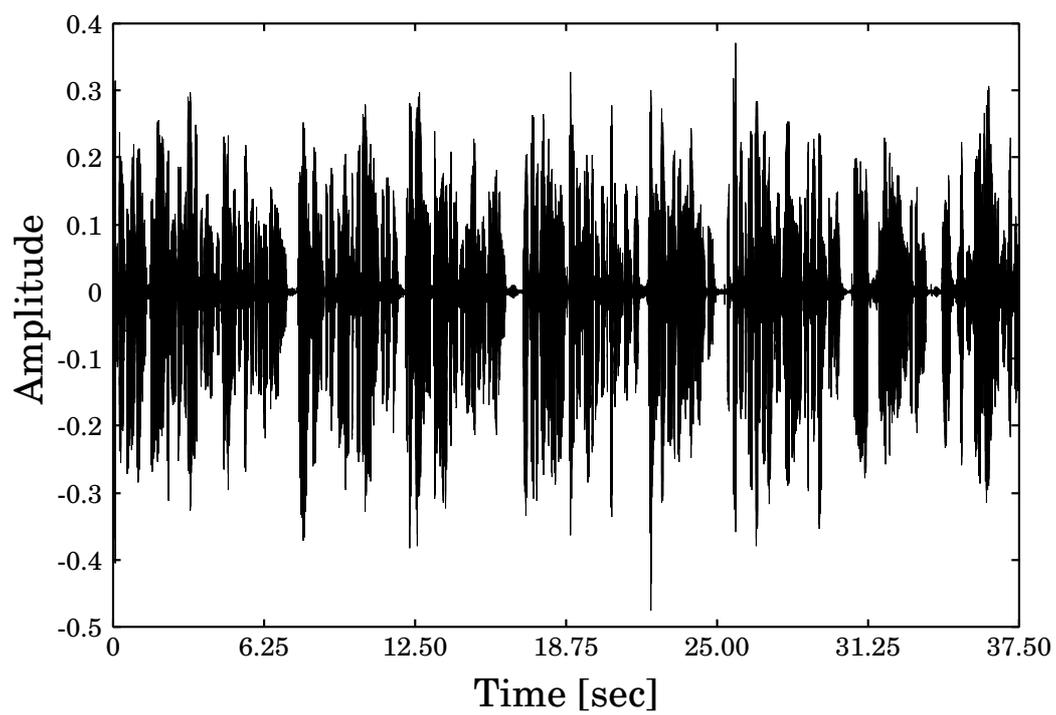


図 3.12 右側原信号

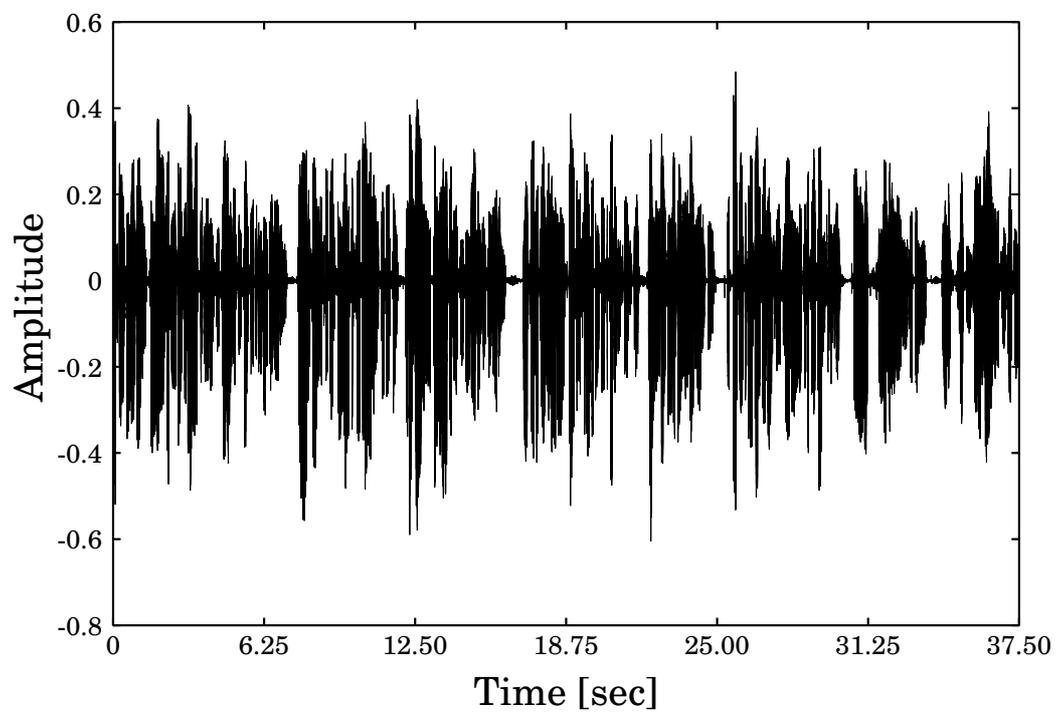


図 3.13 左側原信号

3.5 まとめ

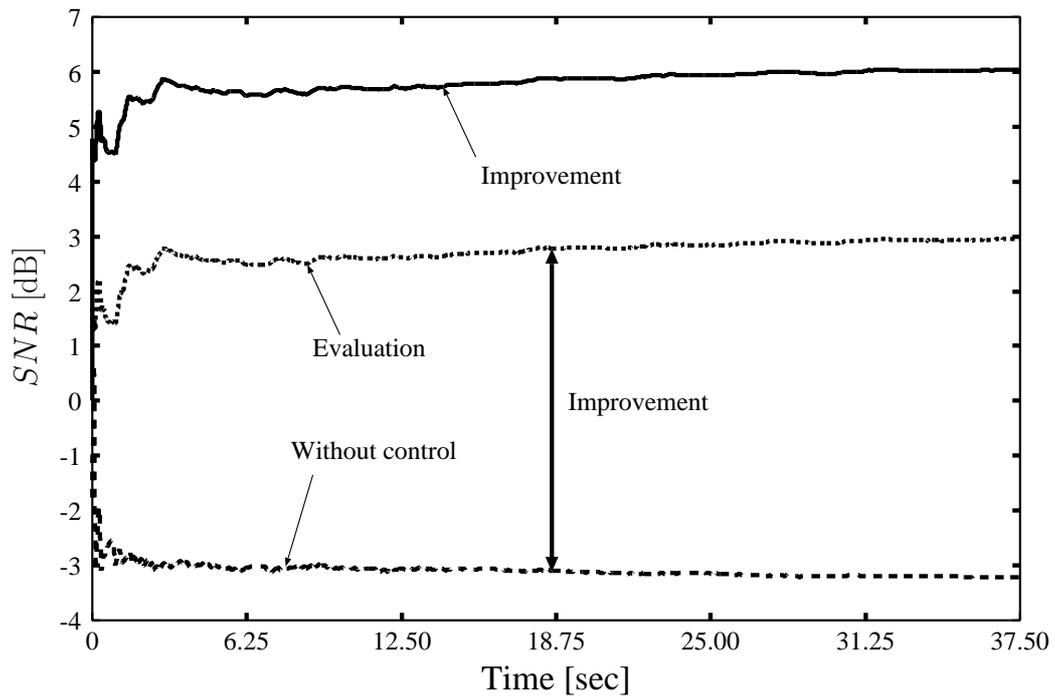


図 3.14 改善量の算出

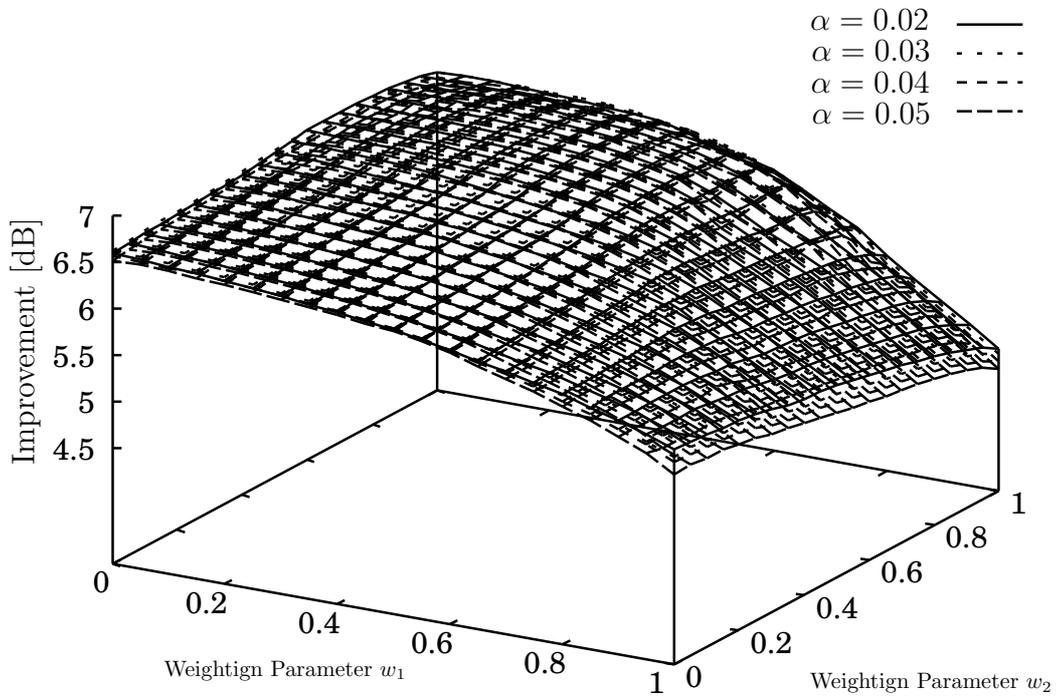


図 3.15 右側改善量に対する収束特性

3.5 まとめ

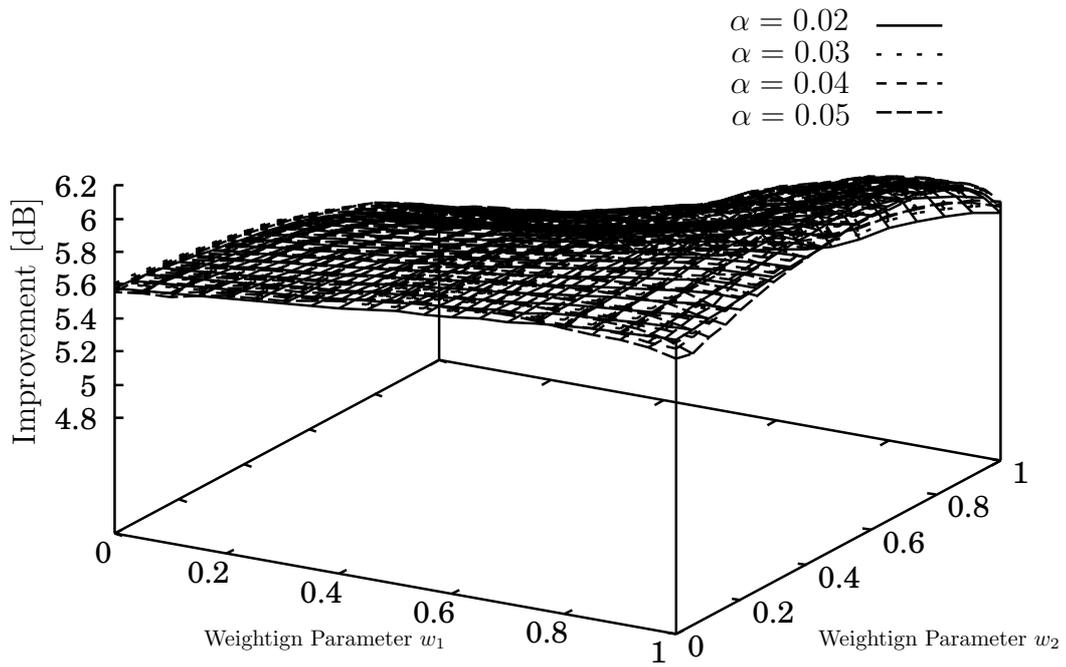


図 3.16 左側改善量に対する収束特性

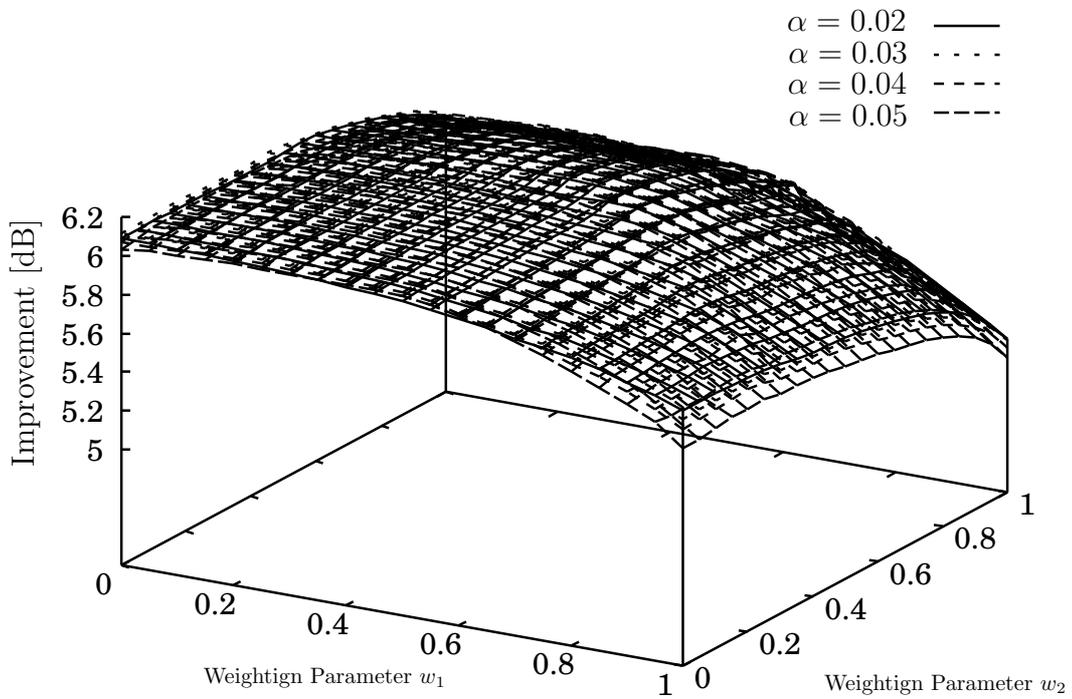


図 3.17 環境 1 に対する収束特性の比較 (左右改善量の平均)

3.5 まとめ

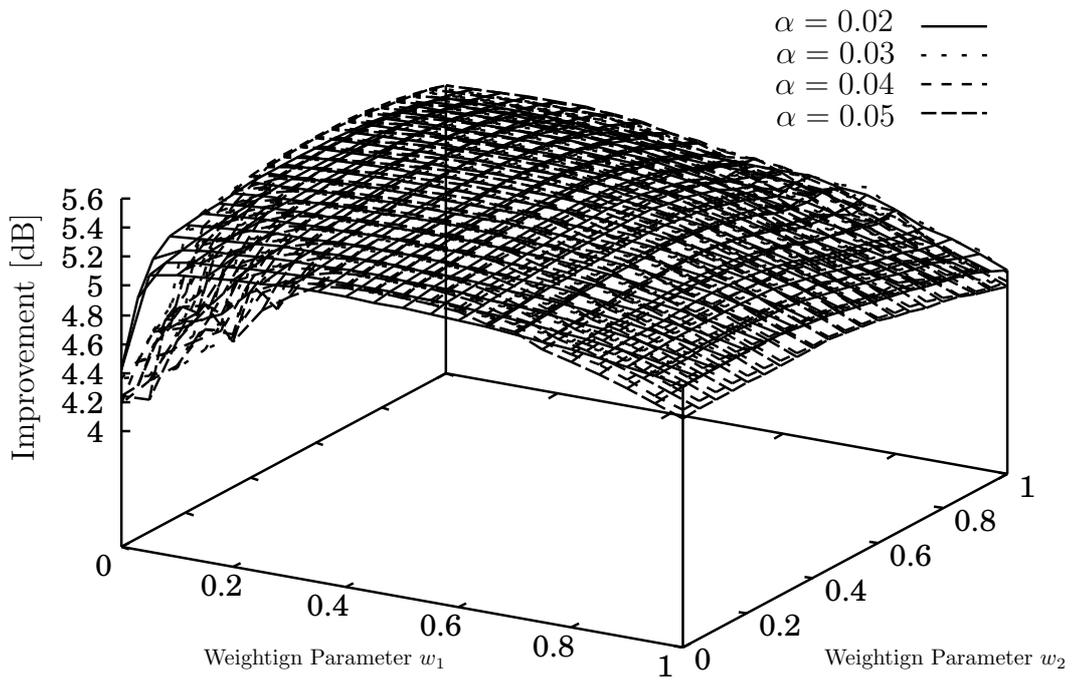


図 3.18 環境 2 に対する収束特性の比較

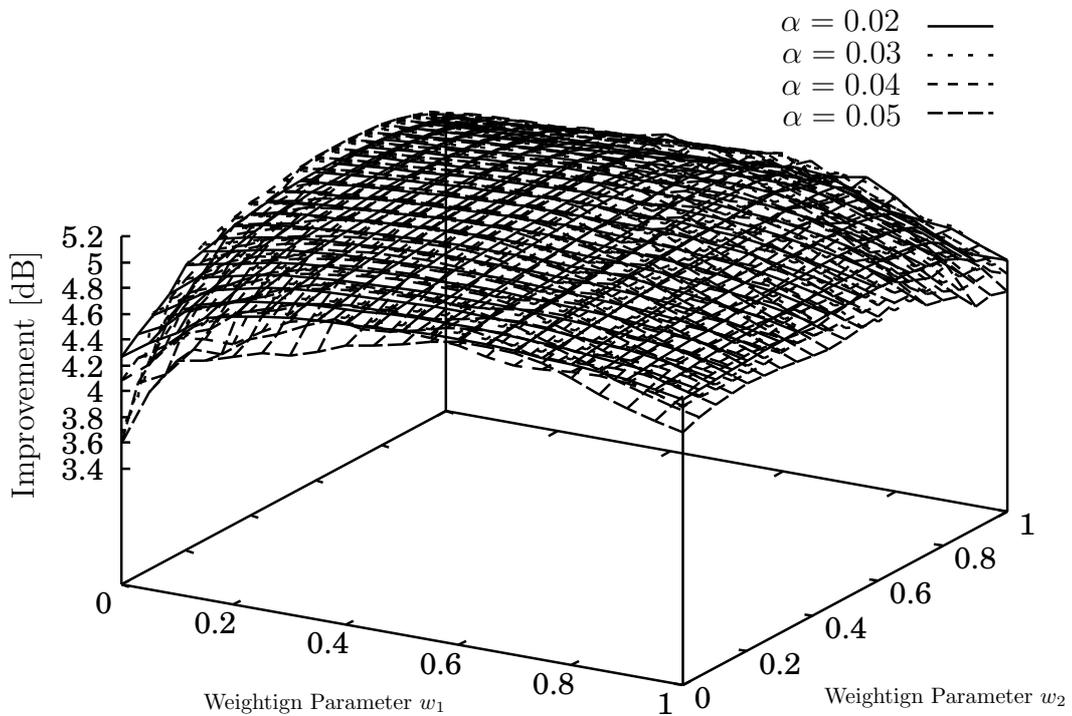


図 3.19 環境 3 に対する収束特性の比較

3.5 まとめ

表 3.2 環境 1 の改善量 ($r_N = 0.73$)

α	w_1	w_2	改善量 [dB]
0.02	0.15	0.05	6.20
0.03	0.10	0.20	6.23
0.04	0.30	0.45	6.21
0.05	0.35	0.25	6.23
0.06	0.25	0.40	6.22
0.07	0.30	0.45	6.20

表 3.3 環境 2 の改善量 ($r_N = 0.55$)

α	w_1	w_2	改善量 [dB]
0.02	0.35	0.50	5.27
0.03	0.35	0.50	5.36
0.04	0.35	0.50	5.40
0.05	0.35	0.50	5.40
0.06	0.30	0.55	5.39
0.07	0.30	0.60	5.38

3.5 まとめ

表 3.4 環境 3 の改善量 ($r_N = 0.45$)

α	w_1	w_2	改善量 [dB]
0.02	0.60	0.75	5.02
0.03	0.50	0.70	5.11
0.04	0.45	0.70	5.13
0.05	0.45	0.85	5.13
0.06	0.55	0.80	5.11
0.07	0.45	0.85	5.09

第 4 章

聴覚特性に基づく出力誤差の算出

4.1 はじめに

音場再生システムでは，伝達系の変動に対する補正フィルタの追従が不可欠であることは 1.2 で述べたとおりである．このことは補正フィルタ係数を算出するために用いる適応フィルタ係数を可能な限り速く収束させることに帰着する．本章では，人間の聴覚特性に基づいた出力誤差の重み付け法を提案し，適応フィルタの収束速度の向上をはかる．

4.2 聴覚特性に基づく出力誤差の算出

前章で述べた多入力信号補正システムにおいて適応フィルタは，伝達特性の逆特性を近似的に推定するために用いられており，この過程で扱った学習同定法は出力誤差が最小となるようにフィルタ係数を更新するアルゴリズムである．通常，この出力誤差に対して余分な周波数帯域の音圧レベルを軽減することは計算効率の向上につながり，フィルタの収束速度は向上する．

ここで，音場再生システムにおける音の再現性は人間が評価することに留意すると，出力誤差を聴覚特性に合わせて重み付けを行うことは，臨場感を損なうことなく，余分な周波数帯域の音圧レベルを軽減できると考えられる．

4.2 聴覚特性に基づく出力誤差の算出

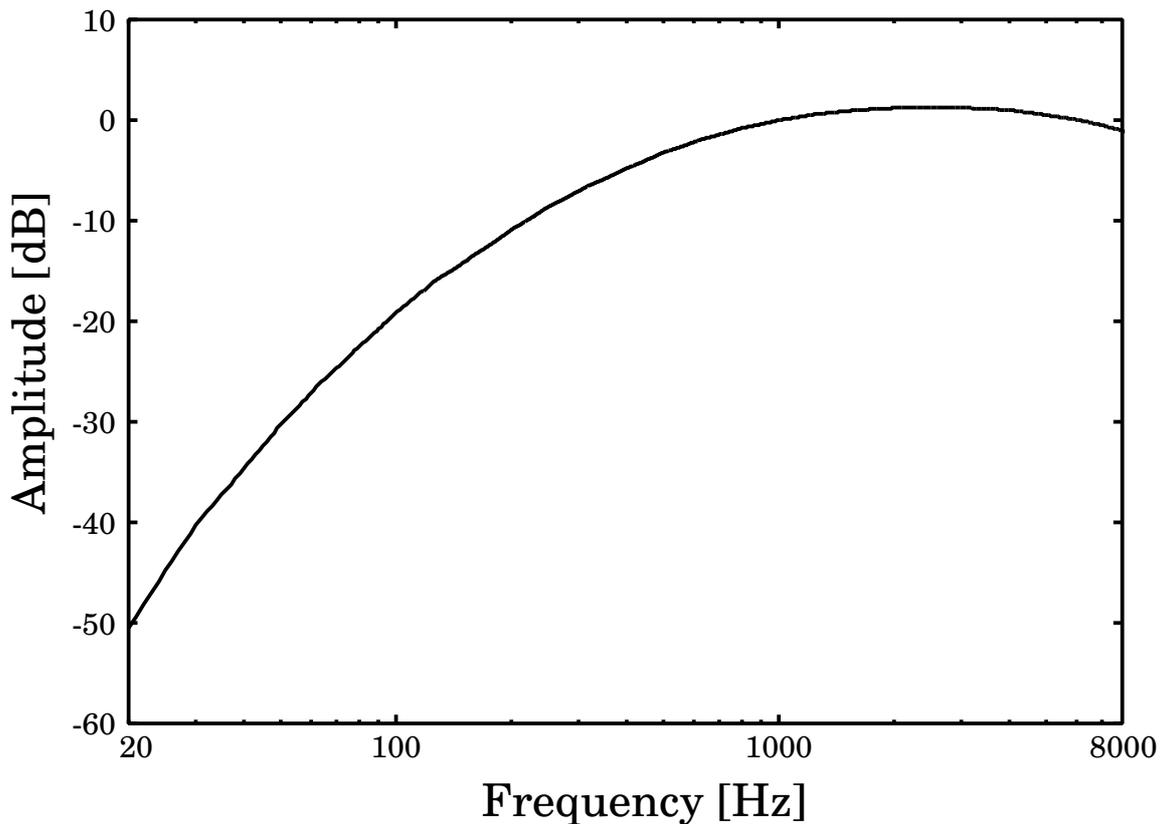


図 4.1 A 特性音圧レベル

4.2.1 重み付け関数の算出

本節では、聴覚特性に基づいた重み付け関数を導出する過程について述べる。人間が感じることのできる音の周波数帯は、約 20Hz~ 20kHz である。この範囲には Robinson らによって測定された等ラウドネス曲線に見られる聴覚特性が存在することが知られている。本論文では、この特性に基づき JIS C 1502 で騒音測定の指標として規格化されている A 特性を用いて、重み付け関数を導出する。図 4.1 に A 特性音圧レベルを示す。これは人間の聴覚の周波数特性を 20Hz~ 8kHz の範囲で模したもので、低周波領域と高周波領域での感度が鈍くなる特性をもっている [19]。この特性から重み付け関数 $\phi(k)$ を次の式により算出する。

$$\phi(k) = \frac{Res(k) + |\min Res|}{|\min Res|} \quad (4.1)$$

4.2 聴覚特性に基づく出力誤差の算出

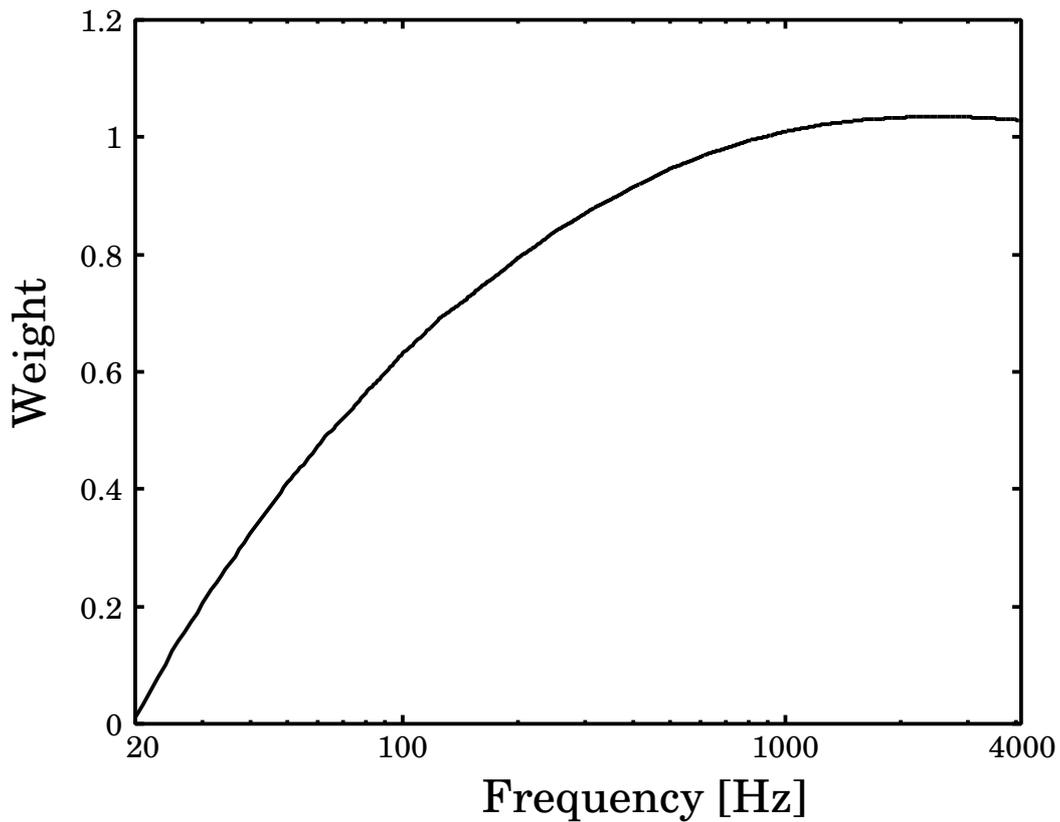


図 4.2 重み付け関数

ここで $Res(k)$, $\min Res$ はそれぞれ A 特性における周波数応答およびその中の最小値を示す . ただし $k < 20$ のとき $\phi(k) = 0$ とする . 図 4.2 に 20Hz~ 4kHz の重み付け関数を示す .

4.2.2 重み付け誤差の算出とフィルタ係数の更新手順

本節では , 適応フィルタに対する重み付け出力誤差の算出方法とフィルタ係数の更新手順について述べる .

図 4.3 に重み付け出力誤差を得るためのブロック図を示す . 3.3.2 で提案した多入力信号補正システムにおける入力信号 $d_i(t)$ と適応フィルタからの出力信号 $s_{ij}(t)$ をそれぞれ周波数領域に変換し , 式 (4.1) によって得られた重み付け関数を用いて , 以下の式でそれぞれの信号に対し重み付けを行う .

$$D'_i(k) = \phi(k)D_i(k) \quad (4.2)$$

4.2 聴覚特性に基づく出力誤差の算出

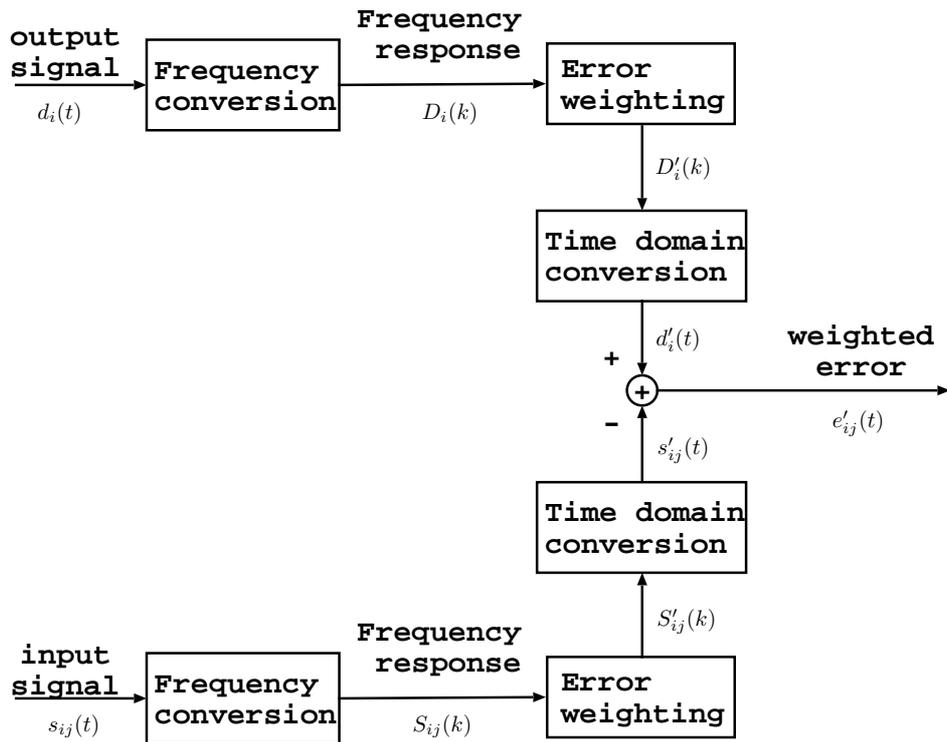


図 4.3 重み付け関数

$$S'_{ij}(k) = \phi(k)S_{ij}(k) \quad (4.3)$$

重み付けされた出力信号 $S'_{ij}(k)$ と入力信号 $D'_i(k)$ をそれぞれ時間領域に変換し，

$$e'_{ij} = d'_i(t) - s'_{ij}(t) \quad (4.4)$$

として新たな重み付け出力誤差を得る．多入力信号補正システムにおいて，重みつき出力誤差を用いた更新手順は

手順 1: 重み付き誤差の算出

$$D_i(k) = \text{FFT}[d_i(t)]$$

$$S_{ij}(k) = \text{FFT}[s_{ij}(t)]$$

$$d'_i(t) = \text{IFFT}[\phi(k)D_i(k)]$$

$$s'_{ij}(t) = \text{IFFT}[\phi(k)S_{ij}(k)]$$

$$e'_{ij}(t) = d'_i(t) - s'_{ij}(t)$$

4.3 計算機シミュレーションによる評価

手順 2: 適応フィルタの係数ベクトルを学習同定法によって算出

$$\mathbf{h}_{ij,L}(t+1) = \mathbf{h}_{ij,L}(t) + \alpha \frac{\mathbf{y}_{ij,L}(t)}{\|\mathbf{y}_{ij,L}(t)\|^2} e'_{ij}(t)$$

手順 3: 左右の所望信号に対する補正フィルタ係数ベクトルの更新

$$\mathbf{c}_{1,L}(t+1) = w_1 \mathbf{h}_{11,L}(t) + (1 - w_1) \mathbf{h}_{21,L}(t)$$

$$\mathbf{c}_{2,L}(t+1) = w_2 \mathbf{h}_{22,L}(t) + (1 - w_2) \mathbf{h}_{12,L}(t)$$

となる．ここで， $\text{FFT}[\cdot]$ ， $\text{IFFT}[\cdot]$ はそれぞれフーリエ変換および，フーリエ逆変換によって得られる値を表す．通常，周波数帯域で重み付けを行うには，演算量を考慮すると聴覚特性をもつバンドパスフィルタを実装することが望ましいが，本論文では等価的にフーリエ変換を用いた聴覚特性の重み付けを行った．

4.3 計算機シミュレーションによる評価

本節では，重み付け出力誤差を用いた多入力信号補正システムの有効性を示すために計算機シミュレーションを行った．

4.3.1 シミュレーション条件

原音である所望信号は bf 3.4.2 と同様のものを用いる．また，適応フィルタおよび補正フィルタのインパルス応答長は $L = 512$ ，再生空間の伝達関数として 3.4.1 の環境 1 で測定した応答長 $L_g = 512$ のインパルス応答を用いる．学習同定法のステップゲインは $\alpha = 0.05$ ，左右の補正フィルタに対する重み付けパラメータはそれぞれ $w_1 = 0.35$ ， $w_2 = 0.25$ とする．

原音に対する再現音の再現精度は，次のような SNR を用いた．

$$SNR \text{ [dB]} = 10 \log_{10} \frac{\sum [x_i(t)]^2}{\sum [y_j(t) - x_i(t)]^2} \quad (4.5)$$

ここで $x_i(t)$ は所望信号の音圧， $y_j(t)$ は観測信号の音圧を表す．ただし，評価量には左右の所望信号に対して一切補正を行わずに算出した SNR と提案手法により算出した SNR との差を改善量として評価を行う．

4.4 まとめ

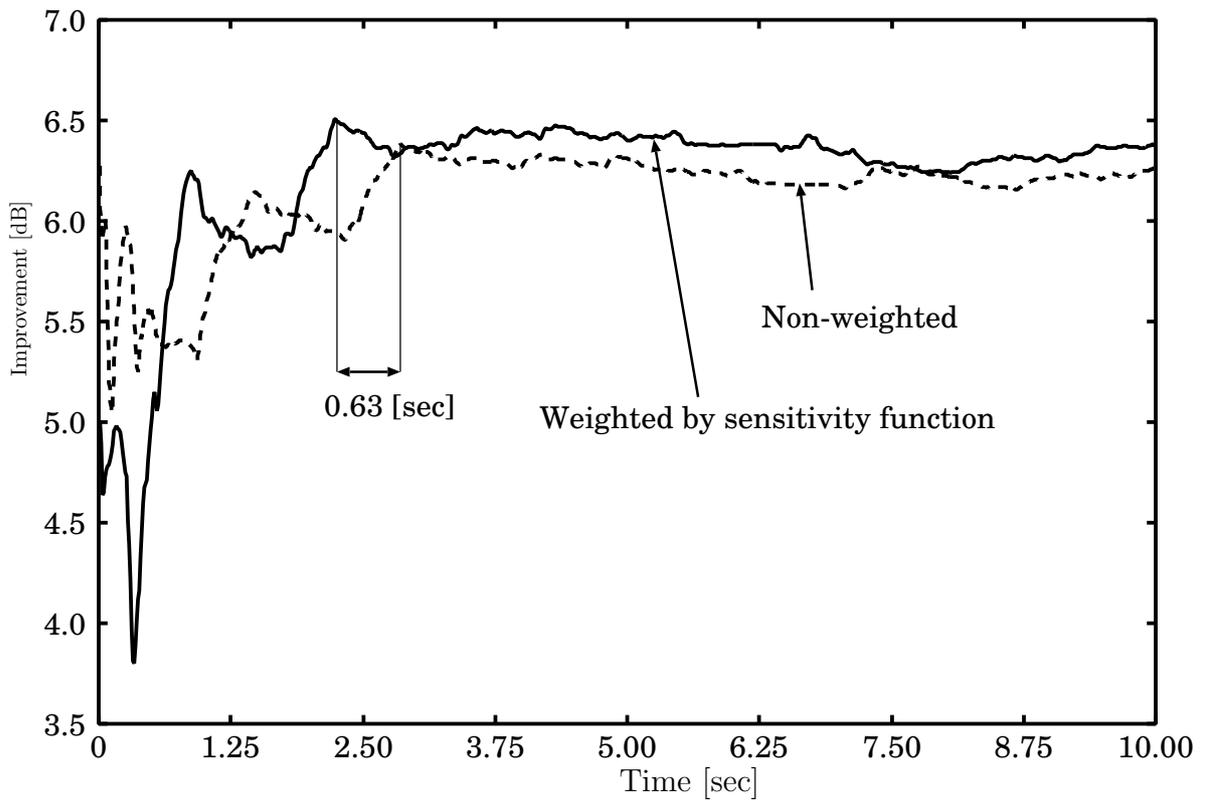


図 4.4 右チャンネルの収束特性の比較

4.3.2 シミュレーション結果

4.3.1 の条件でシミュレーションを行ったときの右チャンネルおよび左チャンネルに対する収束特性について、重み付けを行わない場合と比較したものをそれぞれ図 4.4, 4.5 に示す。これらの結果から、出力誤差算出の際に聴覚特性の重み付けを行ったことで 0.63 秒程度、収束速度の向上していることがわかる。

4.4 まとめ

本章では、多入力信号補正システムにおいて音の再現精度に対する収束速度を向上させるために人間の聴覚特性を利用して、計算効率を高める手法を提案した。計算機シミュレーションを行った結果、左右チャンネルにおいて両者共に、収束速度が向上していることを確認した。これは、人間にとって感度の鈍い周波数帯域に対する処理が抑制され、計算効率に向

4.4 まとめ

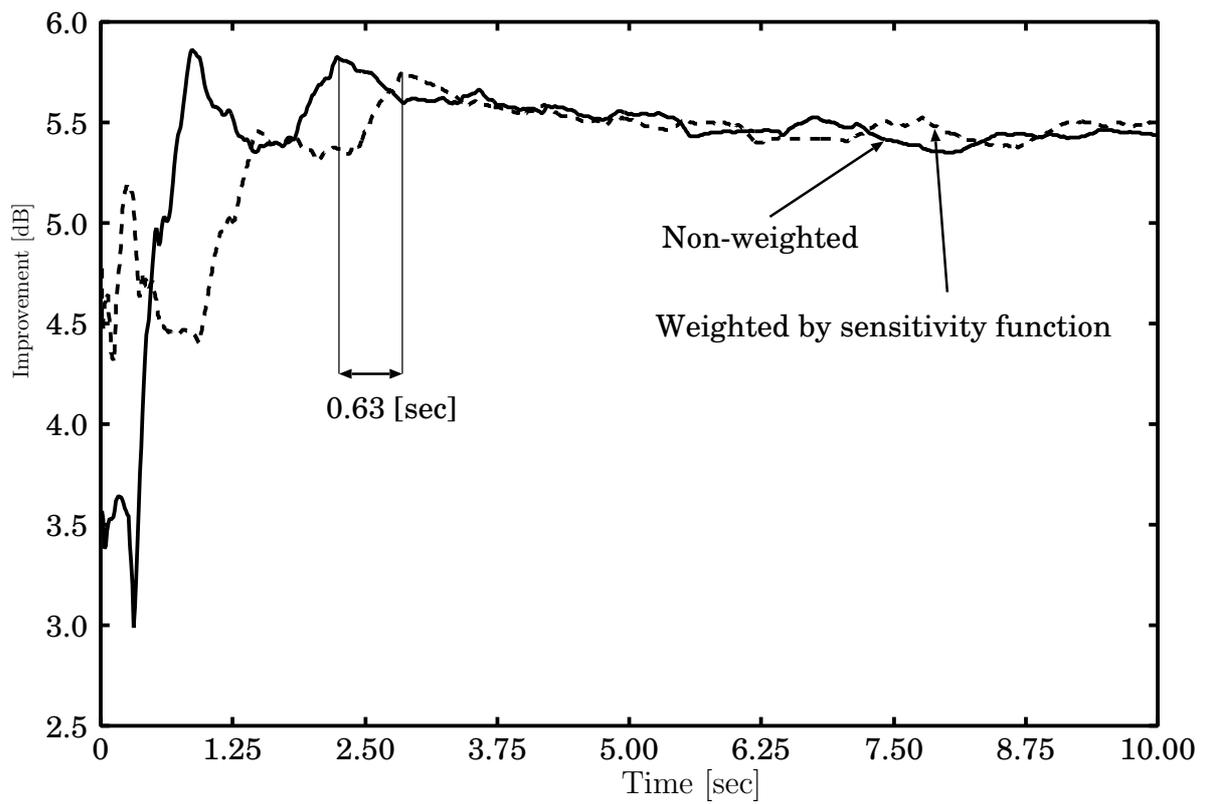


図 4.5 左チャンネルの収束特性の比較

上したためであるといえる。

第 5 章

結言

5.1 本研究のまとめ

ある空間の音場そのものを全く別の空間で再現することは，従来手法では成し得ることのできない夢の技術である．ラウドスピーカ再生によって音場再生を実現する場合，再生する室内の反射や残響を取り除かなければならず，そのためには逆フィルタの設計が不可欠である．しかしながら，逆フィルタが正確な室内伝達特性の逆特性を持つことは現実的には困難であり，それゆえに計算量を抑えた近似逆フィルタの設計手法を確立することが必要である．この問題を解決すべく，本論文では 2 チャネル音場再生システムにおいて新しい制御系を提案し，その効果について検証した．

第 3 章では，クロストークの影響を軽減させるために 2 本の伝達経路における影響を単一の制御系で補正するシステムを提案した．再現精度としては不十分であったものの，クロストーク成分の相互相関が大きいほど，無制御の状態と比べて観測信号が原音に近似している様子が計算機シミュレーションにより確認された．

第 4 章では，伝達特性に追従させるための一提案として，補正フィルタの収束速度を向上させる手法を提示した．計算機シミュレーションから約 0.63 秒程度の向上が確認できた．

5.2 今後の課題

今後の課題として，まずは再現精度を向上させるために何らかの工夫が必要である．また，提案した多入力信号補正システムに用いる重み付けパラメータの最適な設定値につい

5.2 今後の課題

て未検討である．より多くの環境モデルを想定し，実環境データを用いた計算機シミュレーションから，伝達特性と重み付けパラメータとの関連を明らかにする必要がある．

謝辞

本論文は、筆者が高知工科大学大学院情報システム工学コースに在籍中の研究成果をまとめたものです。本研究を遂行するにあたりましては、数多くの方々に大変御世話になりました。

高知工科大学情報システム工学科の福本昌弘助教授には筆者が同大学院に在籍している間、主指導教員並びに論文主査として終始ご指導いただきました。福本助教授のご指導は厳しいご指摘の中にも温かいご助言もあり、福本助教授なくして、本研究の遂行は成し得ませんでした。ここに熱く御礼申し上げます。また、本研究論文の審議ならびに本研究にご助言を下さいました、島村和典教授、坂本明雄教授に深く感謝いたします。さらに、日頃から有益な情報提供や他面にわたる御助力をして頂いた妻鳥貴彦講師に深く感謝致します。

筆者が同大学院に進学して2年間、公私にかかわらず、いつも側で温かく見守ってくれた福本研究室の佐伯幸郎氏には頭が上がりません。佐伯氏なくして本研究の遂行は、まずありえなかったことでしょう。また研究活動以外にも、実践的確率論についてご教授頂いたことは大きく私の人生を変えてくれました。おかげさまでお金の大切さをしみじみと痛感致しました。坂本研究室の河内友彦氏には、他大学から進学して参りました筆者を最初に仲間として温かく向かい入れて頂き、その後、就職活動、セミナーならびに研究活動を共にしてくれたことは、とても心強いものでした。ここに深く感謝致します。TA活動および日頃の研究活動を共にしました、福本研究室の福富英次氏にも深く感謝致します。特に理論武装した日頃の語り口調にはいつも目を見張るものがありとても勉強になりました。博士課程に進学して、福本助教授ご指導の下、すばらしい研究成果を期待しております。

秘書室の村上加織様には、同大学院に進学するにあたりまして、大変お世話になりました。すばらしい先生方やすばらしい仲間めぐり会うことができたのも、村上様の進学のすすめなくしてありえませんでした。本当にありがとうございました。また、加織様のご相談役にいつも力をそそがれておりました、篠森研究室の平山正治氏には授業のことやTAのこ

謝辞

と等でたくさんのアドバイスを頂きました。深く感謝致します。

最後に、個々にはお名前を申し上げられませんが、その他の先輩、同期、後輩の皆様にもいたるところで助けていただきました。本当にありがとうございました。

あらためまして、公私関わらず筆者を支えて下さった全ての方々に御礼申し上げます。

参考文献

- [1] NHK 受信技術センター，知っておきたい地上デジタル放送，日本放送出版協会，2003．
- [2] 大賀寿郎，山崎芳男，金田豊，音響システムとデジタル処理，電子情報通信学会，1995.
- [3] P.A.Nelson, H.Hamada and S.J.Elliott, “Inverse Filters for Multi-Channel Sound Reproduction,” IECE Trans. Fundamentals, vol.E75-A, no.11, pp.1468–1473, Nov.1992.
- [4] 北脇信彦，デジタル音声・オーディオ技術，電気通信協会，1999．
- [5] 辻井重男，久保田一，古川利博，趙晋輝，適応信号処理，昭晃堂，1995.
- [6] 辻井重男，鎌田一雄，デジタル信号処理，昭晃堂，1997．
- [7] 臼井支朗，信号解析，オーム社，1998．
- [8] 電子情報通信学会，デジタル信号処理ハンドブック，オーム社，1993．
- [9] 酒井英昭，信号処理，オーム社，1998．
- [10] 飯國洋二，適応信号アルゴリズム，培風館，2000．
- [11] S. ヘイキン (武部幹 訳)，適応フィルタ入門，現代工学社，1994．
- [12] 久保田一，大石邦夫，C 言語によるデジタル信号処理入門，コロナ社，1999．
- [13] 春日正男，船田哲男，林伸二，武田一哉，音声情報処理，コロナ社，2001.
- [14] 秋山由佳，福本昌弘，“多入力信号補正システムにおけるフィルタ構成法，” デジタル信号処理シンポジウム C1-6，2001 年 11 月．
- [15] 秋山由佳，福本昌弘，“聴覚重み付けによる多入力信号補正システム，” デジタル信号処理シンポジウム A7-4，2002 年 11 月．
- [16] 浜崎 真二，福本 昌弘，“多入力信号補正によるステレオ型音場再生システム，” 第 19 回信号処理シンポジウム講演論文集，A4-3，Nov.2004.
- [17] 浜崎 真二，福本 昌弘，“クロストーク成分における相互相関に着目した音場再生システム，” 信学技報 SIP2004-115，vol.104，no.559，pp.31–36，2005.

参考文献

- [18] J.Nagumo and A.Noda,“A Learning Method for System Identification,” IEEE Trans. AC, vol.12, no.3, pp.282–287, 1967.
- [19] JIS C 1502-1990 , 普通騒音計 , 日本工業規格 , 1997
- [20] JIS C 1505-1988 , 精密騒音計 , 日本工業規格 , 1997

研究業績

国際会議

1. Shinji Hamasaki, Masahiro Fukumoto, “Stereo Type Sound Field Reproduction System using Multi-Input Correction,” International Conference Next Era Information Networking (NEINE'04), pp.468–475, Kochi, Japan, Sep. 2004.

研究会・大会発表

1. 浜崎真二，河合一慶，宮内新，荒井秀一，“簡単な CPU を用いたハードウェア教育” FIT2002，N-26，2002.
2. 浜崎真二，福本昌弘，“多入力信号補正によるステレオ型音場再生システム，” 電気関係学会四国支部連合大会，p.15，2003.
3. 浜崎 真二，福本 昌弘，“多入力信号補正によるステレオ型音場再生システム，” 第 19 回信号処理シンポジウム講演論文集，A4-3，Nov.2004.
4. 浜崎真二，福本 昌弘，“クロストーク成分における相互相関に着目した音場再生システム，” 信学技報 SIP2004-115，vol.104，no.559，pp.31–36，2005.