

平成 16 年度
学士学位論文

多入力信号補正における収束特性の評価法

Evaluation of settling characterisitic by multiinput
signal correction

1050348 豊永 雅行

指導教員 福本 昌弘

2005 年 3 月 11 日

高知工科大学 情報システム工学科

要 旨

多入力信号補正における収束特性の評価法

豊永 雅行

原音場の特性を再生音場にて忠実に再現する手法として、多チャネル-多点制御系は有効である。しかしながら、再生空間の残響時間が長い環境では制御系が複雑になり、その結果演算量が膨大になるため実用化にいたっていない。このことに対し、浜崎らは空間伝達特性の影響を軽減させる2チャネル-2点制御系（多入力信号補正システム）を提案してきたが、伝達特性の近似逆特性の収束性については保証されていない。そこで、本論文では様々な実環境を想定し、それぞれの環境における伝達関数の実測値を用いて計算機シミュレーションを行い、重み付けパラメータやクロストークの相互相関係数を用いて収束特性の検証を行う。その結果から相互相関係数が高い程システムの精度がよくなることを確認している。また、重み付けパラメータの設定には所望信号のパワー比に関係あることも確認している。このことから、クロストーク成分の相互相関係数が高い程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 キーワード:多入力信号補正システム, 伝達特性

Abstract

Evaluation of settling characterisitic by multiinput signal correction

Masayuki Toyonaga

The multi channel - multipoint control system, technique for faithfully reproducing characteristic of original sound place in reproduction sound place is effective. But, in the environment with long reverberation time of the reproduction space because the control system becomes complex, and the amount of the operation becomes huge as a result, it has not become practical use. The two channel - two point control system “multiinput signal correction system” have been proposed by Hamasaki to reduced the influence of the space transmission characteristic. And, settling to approximation reverse-characteristics of the transmission characteristics are not guaranteed.

In this resarch, the transmission characteristic is verified by simulating the computer by assuming variety of real environmental models, the setting characterisitic is c]verified by using the weight parameter and cross correlation coefficient. As results, we have showed that the accuracy of the system improves by the cross correlation coefficient high. And, we have showed confirmed to the setting of the weight putting parameter to relate compared with the power of the desire signal. Moreover, we have showed the tendency to reproduce sound field by computer simulateion by raising the cross correlation correlarion coefficient of crosstalk.

key words keyword: Multiinput signal correction, Transmission characterisitic

目次

第 1 章	緒論	1
1.1	本研究の背景	1
1.2	先行研究について	1
1.3	本研究の目的	2
1.4	本論文の構成	2
第 2 章	音響信号処理モデル	4
2.1	まえがき	4
2.2	FIR デジタルフィルタ	4
2.3	適応信号処理	6
2.3.1	適応フィルタ	6
2.3.2	適応アルゴリズム	8
2.3.3	学習同定法	8
2.4	まとめ	10
第 3 章	音場制御	11
3.1	まえがき	11
3.1.1	トランスオーラルシステム	12
3.2	多チャネル-多点制御系	14
3.3	2 チャネル-2 点制御系	16
3.4	それぞれの問題点	17
3.5	まとめ	18
第 4 章	多入力信号補正システム	19
4.1	まえがき	19

目次

4.2	多入力信号補正システム	19
4.3	まとめ	21
第 5 章	計算機シミュレーション	22
5.1	まえがき	22
5.2	シミュレーション条件	22
5.3	シミュレーション結果	24
5.4	まとめ	25
第 6 章	結論	29
6.1	本研究における評価	29
6.2	今後の課題	29
	謝辞	30
	参考文献	31

図目次

2.1	FIR デジタルフィルタ	5
2.2	適応フィルタを用いたブロック図	7
3.1	トランスオーラルシステムの模式図	12
3.2	室内伝達特性の補正	13
3.3	多チャネル-多点制御	15
3.4	MINT の考え方に基づいた多チャネル-多点制御の構成	15
3.5	2チャネル-2点制御	16
3.6	2チャネル-2点制御系のブロック図	17
4.1	多入力信号補正システム	20
5.1	改善量の算出	23
5.2	パターン 1	25
5.3	パターン 2	25
5.4	パターン 3	26
5.5	パターン 4	26
5.6	パターン 5	26
5.7	パターン 6	27
5.8	パターン 7	27

表目次

2.1 図 2.2 における変数の説明	7
5.1 チャネル 1 のシミュレーション結果	28
5.2 チャネル 2 のシミュレーション結果	28

第 1 章

緒論

1.1 本研究の背景

デジタル技術の著しい進展によって、映像や音声など様々なデータの伝送が簡単に可能となった。その中でも、音場再生において部屋にいても映画館やコンサートホールにいるような立体感を感じることが従来の技術で可能となっている。例えば、5.1ch サラウンドシステム等がある。しかし、この技術は実際の音源を再生することで臨場感を与えるのではなく物理的に音を周りから発することであたかもその場所に居るように聞こえるように作用している。この方法では、スピーカの数が多くなり、コストの増加が考えられる上に、部屋のスペースなどを考慮しなければならない。

一方、実際の音源を再生する方法としては、原音場でマイクロフォンを用いて音を収録し、再生音場ではスピーカで音を再生している。ただし、原音場および再生音場での伝達特性を考慮する必要がある。その解決方法としていくつかの方法が提案されているが、再生音場に制限があることや、残響が多かったり再生音場に制限があることや受聴者は時間や場所に拘束されてしまう等、技術的にたくさんの問題点があるため実現が困難であると考えられている。

1.2 先行研究について

スピーカを用いて音場再生が行われた場合、1つのスピーカから発せられた音は、左右耳に対して2つの異なる経路で伝達されるため、通常2つの補正フィルタを必要とする。しか

1.3 本研究の目的

し、1つのスピーカは1つの補正フィルタに対する処理しか行うことができない。そこで、秋山らによって1つのフィルタにより2つの伝達経路を制御する信号補正法が提案されている。また、それを元に浜崎らによってステレオ型の音場再生システム(多入力信号補正システム)が提案されている。このシステムでは、1つの受聴点に到達する2本の伝達特性の影響を除去するように補正をかけることでクロストークの影響が軽減できる。また、フィルタの数も通常のスピーカより少ないため演算量の軽減もなされている。しかし、クロストークによる問題からフィルタの収束が数学的には解けないため、適応フィルタの収束性が未保証となっている。

1.3 本研究の目的

本研究では、多入力信号補正システムを使用して様々な実環境における伝達特性の実測値を用いて適応フィルタの収束特性の検証を行う。また、補正フィルタのパラメータ更新に用いる重み付けパラメータの最適な値が求まっていないため、最適な重み付けパラメータの決定方法を検討する。

1.4 本論文の構成

ここで、本論文の構成について述べる。

第2章では、デジタル信号処理の基礎技術である適応信号処理について述べる。信号処理を行う際に用いられる適応アルゴリズムのうち、今回使用している多入力信号補正システムで使われている学習同定法についての説明を行う。

第3章では、音場制御としてスピーカを用いて構築するトランスオーラルシステムについて説明する。また、トランスオーラルシステムをMINT法に基づいて構築された多チャネル-多点制御系とそれを改良した2チャネル-2点制御系について述べる。

第4章では、2チャネル-2点制御系である多入力信号補正システムに付いて述べる。

第5章では、第4章で示す多入力信号補正システムの収束特性の検証を様々な実環境モデ

1.4 本論文の構成

ルを想定したそれぞれの環境における伝達特性の実測値を用いて計算機シミュレーションにより行う。また、補正フィルタパラメータの更新に必要な重み付けパラメータの最適な値の決定法も検討する。

最後に、第6章では、本研究での計算機シミュレーション結果を評価し、問題点を指摘することで今後をの課題を考える。

第 2 章

音響信号処理モデル

2.1 まえがき

音楽、音声などの音響信号は連続的な値を持つアナログ信号である。このアナログ信号を離散的な値を持つデジタル信号に変換することで、コンピュータ上でデータの処理を行うことができる。

処理の対象となる原音場で出力されて受聴点に達するまでに所望信号に伝達関数が付加した観測信号は、入力されてから観測されるまでの間のシステムに関する何らかの情報を保持している。そこで、観測信号から伝達関数の逆特性を近似する信号成分を抽出する機能を持ったデジタルフィルタについて述べる。

デジタルフィルタでは、入力にある処理を行い変形することによって、必要とする出力を得る。デジタルフィルタは大きく分けて、有限長の応答を持つ FIR(Finite Impulse Response) フィルタと、無限長の応答をもつ IIR(Infinite Impulse Response) フィルタに分類される。本章では、本研究で用いる多入力信号補正システムに組み込まれている FIR フィルタについて述べた後、適応フィルタと適応アルゴリズムについて説明を行う。さらに適応アルゴリズムの一つである学習同定法について述べる。

2.2 FIR デジタルフィルタ

デジタルシステムでは、ある信号に対して個々の値を一時的に記憶するため、過去の値を取り出すことが可能である。そこで図 2.1 のように単位時間分の遅延を示す z^{-1} を有限個

2.2 FIR デジタルフィルタ

用いたフィルタを考える.

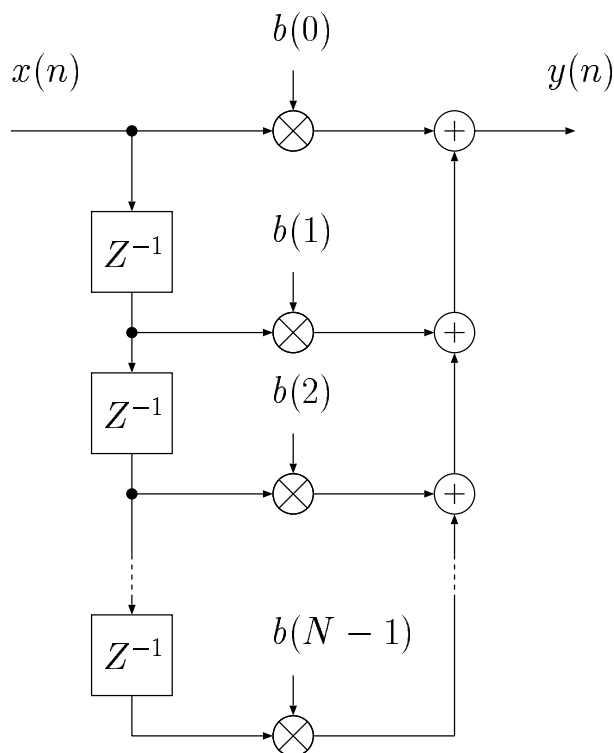


図 2.1 FIR デジタルフィルタ

図 2.1 では, 入力信号 $x(k)$ が各タップにおいてパラメータ (フィルタ係数) と乗ぜられ, それらが加算された結果 $y(k)$ が出力される. このフィルタの入出力関係は,

$$y(k) = \sum_{i=0}^{N-1} b(i)v(k-i+1) \quad (2.1)$$

で与ええられる.

このように, インパルス応答の長さが有限長となるようなデジタルフィルタのことを FIR(Finite Impulse Response) デジタルフィルタと呼ぶ.

式 (2.1) から, 出力信号 $y(k)$ はパラメータ b_n に依存することがわかる. このことは, 同一の入力信号を与えた場合でもパラメータが異なると得られる結果も異なることを表している. ただし, パラメータ w_n とは,

$$b_n = [b(0), b(1), \dots, b(N)]^T \quad (2.2)$$

2.3 適応信号処理

で表される．ここで A^T は A の転置を表す．

2.3 適応信号処理

信号処理では、観測された信号にフィルタを作用させることによって、望ましい信号が出力されるように処理を行う．このとき、フィルタの特性を決めるパラメータが固定である場合は、時間の変化に関係なく処理を行うことが可能である．しかし、観測信号の統計的性質が時間と共に変化する場合には、固定的な処理では正確に対応できない．そこで、変化する信号の性質に応じて、パラメータを更新することのできる適応フィルタを用いることにより対応している．このような処理を行うために、Widrow の先駆的な研究により始められたのが適応信号処理 [3] である．

2.3.1 適応フィルタ

次に、FIR 形で構成された未知システムのパラメータ (インパルス応答) を推定することを考える．図 2.2 にシステムのブロック図を示す．ただし、図 2.2 における変数についての説明は表 2.1 に示す．

このとき、入力信号 $s(t)$ と受聴点におけるの出力信号は $y(t)$ はそれぞれ

$$s(t) = \sum_{i=0}^{N-1} b(i)v(t-i+1) \quad (2.3)$$

$$y(t) = \sum_{i=0}^{N-1} h(i)v(t-i+1) \quad (2.4)$$

と表すことが出来る．また出力誤差は $e(t)$ は

$$e(t) = s(t) - y(t) \quad (2.5)$$

と表される．

式 (2.3), (2.4) において

$$b(i) = h(i) \quad ; \forall i = 0, 1, \dots, N-1 \quad (2.6)$$

2.3 適応信号処理

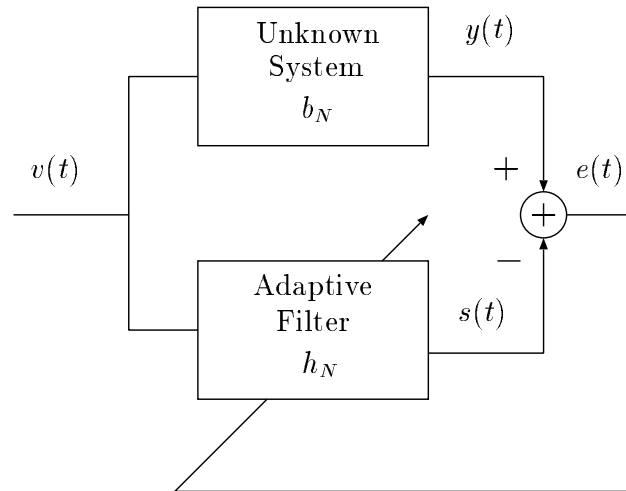


図 2.2 適応フィルタを用いたブロック図

であれば，常に同じ出力を得ることができる．したがって，未知システムのインパルス応答長が有限で，その個数が既知であるならば，同一の入力信号に対して完全に等しい出力結果を与える FIR デジタルフィルタを構成することが可能である．

表 2.1 図 2.2 における変数の説明

t	時刻
N	インパルス応答長 (フィルタタップ数)
$v(t)$	入力信号
$\mathbf{b}_N$	未知システムのフィルタ係数 $(\mathbf{w}_N = [b(0), b(1), \dots, b(N)]^T)$
$\mathbf{h}_N$	適応フィルタのフィルタ係数 $(\mathbf{h}_N = [h(0), h(1), \dots, h(N)]^T)$
$s(t)$	未知形の出力信号
$y(t)$	推定系の出力信号
$e(t)$	出力誤差

2.3 適応信号処理

2.3.2 適応アルゴリズム

適応アルゴリズムとは、各時刻で観測される入力信号 $v(t)$ と出力誤差 $e(t)$ を用いて、パラメータを修正することによって最適解を求める計算手順のことである。

適応アルゴリズムは、1960 年に Widrow-Hoff の LMS アルゴリズムにはじまり、1967 年にこれとは独立に、野田と南雲により学習同定法が発表された。また、その他にも RLS アルゴリズム、BLMS アルゴリズムなどが挙げられる。

図 2.2 において、入力信号 $v(t)$ に対する未知系出力 $y(t)$ は、

$$y(t) = \sum_{i=0}^{N-1} h(i)v(t-i+1) \quad (2.7)$$

と表される。また、式 (2.7) は入力ベクトル内積として

$$y(t) = \mathbf{h}_N^T \mathbf{v}_N(t) \quad (2.8)$$

と表すことができる。ただし、 $\mathbf{h}_N, \mathbf{v}_N$ はそれぞれ

$$\mathbf{h}_N = [h(0), h(1), h(2), \dots, h(N-1)]^T \quad (2.9)$$

$$\mathbf{v}_N = [v(1), v(2), v(3), \dots, v(N-1)]^T \quad (2.10)$$

で定義される。

ここで、誤差の 2 乗平均値を評価量 J として

$$\begin{aligned} J &= E[e^2(t)] \\ &= E[(s(t) - y(t))^2] \\ &= E\left[\left(s(t) - \mathbf{h}_N^T \mathbf{v}_N(t)\right)^2\right] \end{aligned} \quad (2.11)$$

とする。ただし、 $E[\cdot]$ は期待値を表す。

2.3.3 学習同定法

学習同定法は、LMS アルゴリズムとは独立に導かれた適応アルゴリズムである。しかし、別名 NLMS(Normalized-LMS) アルゴリズムと呼ばれ、LMS アルゴリズムのパラメータ修

2.3 適応信号処理

正項をフィルタの状態ベクトルで正規化されたものとみなすことができる．ある時刻 k において，推定系出力 $y(t)$ が未知系出力 $s(t)$ に等しいとすると

$$s(t) = \mathbf{h}_N^T \mathbf{v}_N(t) \quad (2.12)$$

と表すことができる．

しかし， $h_N = b_N$ を満たすためには，すべての所望信号 $v(t)$ に対して式 (2.12) が成り立たなければならない．

そこで，式 (2.12) を満たす解集合の代表ベクトルを $h_N(t)$ とする．この解集合は，式 (2.12) より，入力ベクトル $v_N(t)$ に直行しているといえる．更に， w_N はこの解集合に含まれているので， $h_N(t)$ はある点から $v_N(t)$ 方向にパラメータを修正したとき，もっとも b_N に近い点といえる．

したがって， $h_N(t)$ を b_N に更に近づけるためには，適当に定めたある点よりも b_N より近い $h_N(t+1)$ を次ぎの修正パラメータの初期値とすればよい．以上のことより，

$$\begin{aligned} \mathbf{h}_N(t+1) &= \mathbf{h}_N(t) + (\mathbf{h}_N(t+1) - \mathbf{h}_N(t)) \\ &= \mathbf{h}_N(t) + \frac{(\mathbf{b}_N - \mathbf{h}_N(t))^T (\mathbf{h}_N(t+1) - \mathbf{h}_N(t))}{\|\mathbf{h}_N(t+1) - \mathbf{h}_N(t)\|} \\ &\quad \cdot \frac{\mathbf{h}_N(t+1) - \mathbf{h}_N(t)}{\|\mathbf{h}_N(t+1) - \mathbf{h}_N(t)\|} \end{aligned} \quad (2.13)$$

となる．但し， $\|\cdot\|$ はベクトルのユークリッドノルムを表し，要素の 2 乗和の平方根と定義する．また，式 (2.13) において

$$\frac{(\mathbf{b}_N - \mathbf{h}_N(t))^T (\mathbf{h}_N(t+1) - \mathbf{h}_N(t))}{\|\mathbf{h}_N(t+1) - \mathbf{h}_N(t)\|}$$

はパラメータの修正量を示し

$$\frac{\mathbf{h}_N(t+1) - \mathbf{h}_N(t)}{\|\mathbf{h}_N(t+1) - \mathbf{h}_N(t)\|}$$

はパラメータの修正方向を示す．

ここで

$$\frac{\mathbf{h}_N(t+1) - \mathbf{h}_N(t)}{\|\mathbf{h}_N(t+1) - \mathbf{h}_N(t)\|} = \frac{\mathbf{v}_N(t)}{\|\mathbf{v}_N(t)\|} \quad (2.14)$$

2.4 まとめ

$$\begin{aligned}(\mathbf{b}_N - \mathbf{h}_N(t))^T \mathbf{v}_N(t) &= d(t) - y(t) \\ &= e(t)\end{aligned}\tag{2.15}$$

が成立するので，式 (2.13) は

$$\mathbf{h}_N(t+1) = \mathbf{h}_N(t) + \frac{\mathbf{v}_N(t)}{\|\mathbf{v}_N(t)\|^2} e(t)\tag{2.16}$$

と変形できる．

学習同定法では，パラメータ更新式は，式 (2.16) の修正ベクトルにステップゲインを掛け

$$\mathbf{h}_N(t+1) = \mathbf{h}_N(t) + \alpha \frac{\mathbf{v}_N(t)}{\|\mathbf{v}_N(t)\|^2} e(t)\tag{2.17}$$

で与えられる．

2.4 まとめ

本章では多入力信号補正システムにおける補正フィルタのパラメータ推定時に必要となる音響モデルに付いて述べた．また，多入力信号補正システムの適応フィルタに使われている学習同定法について述べた．

第 3 章

音場制御

3.1 まえがき

音場再生を行うにあたって原音を制御点において忠実に再現することが望まれている。一般に，室内での信号の伝達は，音源から直接到達する音（直接音）の他に，壁などで反射された多数の音（反響音，残響）により構成されている。この反響音，残響，空間の状態など音が受聴者に伝わる際に付加する特性を伝達特性と呼ぶ。この伝達特性を計算機によって処理することによって実際の音場を再現しようとするのが音場再生システムである。しかし，信号の伝達特性は，空間の変化や受聴者の移動に伴い常に一定にはならない。また，信号の伝達特性をもつ尾荒メータは未知であり，常に正確に表現することはできない。そのため，音場再生システムを実現するためには，これらの問題を解決する必要がある。

現在，音場再生を実現するためには，バイノーラルシステムやトランスオーラルシステムが既に研究されている。バイノーラルシステムでは受聴者の耳元でヘッドホンを利用することにより忠実に再現している。しかし，これには長時間のヘッドホン装着により受聴者に負担がかかってしまうことや受聴場所が限定されてしまうという問題がある。トランスオーラルシステムではスピーカを用いた音場再生であるため，バイノーラルシステムの問題点を解消することができる。しかし，スピーカを用いた音場再生では，室内伝達特性の影響を受けるため，これを除去するフィルタを構成することが必要となる。通常，2つの伝達特性を補正するためには，2つのフィルタを必要とする。そのため，ステレオ録音，再生を行う場合には4つの伝達経路を補正するために，4つのフィルタが必要になる。そのため，スピーカの数が増え制御系が複雑になってしまうことから計算量が膨大になる問題点がある。

3.1 まえがき

本章ではスピーカーを用いた音場再生方法としてトランスオーラルシステムについて述べ、スピーカーを用いる際のクロストーク問題と補正フィルタの必要性について述べる。さらに、制御点においてできるだけ少ないスピーカーで補正するための方法として MINT 法に基づいた多チャンネル多点制御系とそれを改良した 2 チャンネル 2 点制御系について述べる。

3.1.1 トランスオーラルシステム

トランスオーラルシステムとは、スピーカーを用いて音場再生を実現しようとする技術である。スピーカーを用いた信号再生では、所望信号に室内伝達系の特性が付加された音が観測される。したがって、所望信号と観測信号を同じにするためには室内伝達特性の影響を除去するためのフィルタが必要となる。

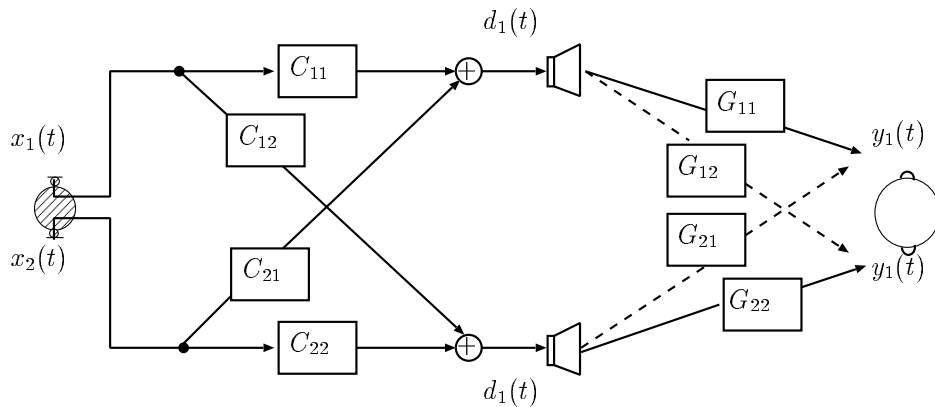


図 3.1 トランスオーラルシステムの模式図

室内伝達特性 $G(z)$ を補正するモデル図 3.2 を示す。音場再生では、所望信号 $x(t)$ を観測信号 $y(t)$ として得ることを目的としている。空間への入力信号 $d(t)$ は、所望信号 $x(t)$ がフィルタ $C(t)$ と通過したものとして与える。また、観測信号 $y(t)$ は、入力信号 $d(t)$ が室内伝達特性 $G(z)$ の影響を受けることにより得られるものとする。ここで、システムが線形系であると仮定すると

$$D(z) = C(z)X(z) \quad (3.1)$$

$$Y(z) = G(z)D(z) \quad (3.2)$$

3.1 まえがき

となる．式 (3.1), (3.2) より

$$Y(z) = G(z)C(z)X(z) \quad (3.3)$$

となり

$$G(z)C(z) = 1 \quad (3.4)$$

となるようにフィルタ $C(z)$ を導入すると所望信号が観測信号の近似となるように設計することができる．このように，室内伝達特性の影響を除去するために構成されたフィルタを補正フィルタと呼ぶこととする．トランスオーラルシステムでは，バイノーラルシステムと同

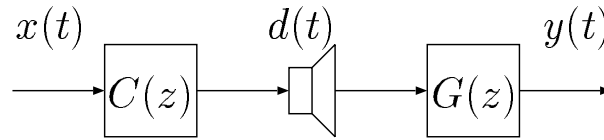


図 3.2 室内伝達特性の補正

様に原音場に設置されたダミーヘッドで 2 チャンネル (ステレオ) 録音を行う．その後，録音された音がスピーカーを用いて再生される．トランスオーラルシステムの模式図を図 3.1 に示す．このとき，観測信号と入力信号の関係は

$$Y_1(z)G_1(z)D_1(z) + G_2(z)D_2(z) \quad (3.5)$$

$$Y_2(z) = G_1(z)D_1(z) + G_2(z)D_2(z) \quad (3.6)$$

となり行列

$$\begin{bmatrix} Y_1(z) \\ Y_2(z) \end{bmatrix} = \begin{bmatrix} G_{11}(z) & G_{12}(z) \\ G_{21}(z) & G_{22}(z) \end{bmatrix} \begin{bmatrix} D_1(z) \\ D_2(z) \end{bmatrix} \quad (3.7)$$

で表すことが出来る．これは，本来は右耳 (左耳) にだけ伝えられるべき音が，反対側の耳にも聞こえるという現状が生じることを意味する．この現象はクロストークと呼ばれ，原音場を再現するために妨害要因となる．そこで，室内伝達特性の影響とクロストークを除去するために補正フィルタを用いられる．所望信号と入力信号の関係を式 (3.7) と同様に行列で

3.2 多チャネル–多点制御系

表すと

$$\begin{bmatrix} D_1(z) \\ D_2(z) \end{bmatrix} = \begin{bmatrix} C_{11}(z) & C_{12}(z) \\ C_{21}(z) & C_{22}(z) \end{bmatrix} \begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} \quad (3.8)$$

となる．室内伝達経路系，補正フィルタ系をそれぞれ

$$\mathbf{G}(z) = \begin{bmatrix} G_{11}(z) & G_{12}(z) \\ G_{21}(z) & G_{22}(z) \end{bmatrix} \quad (3.9)$$

$$\mathbf{C}(z) = \begin{bmatrix} C_{11}(z) & C_{12}(z) \\ C_{21}(z) & C_{22}(z) \end{bmatrix} \quad (3.10)$$

とおく．また

$$\mathbf{Y}(z) = \begin{bmatrix} Y_1(z) \\ Y_2(z) \end{bmatrix} \quad (3.11)$$

$$\mathbf{X}(z) = \begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} \quad (3.12)$$

とおくと式 (3.7)，(3.8) より

$$\mathbf{Y}(z) = \mathbf{G}(z)\mathbf{C}(z)\mathbf{X}(z) \quad (3.13)$$

となり

$$\mathbf{G}(z)\mathbf{C}(z) = \mathbf{I} \quad (3.14)$$

の関係を持つことにより室内伝達特性の逆特性を求める補正フィルタを構成する．

3.2 多チャネル–多点制御系

多チャネル–多点制御系とは原音場でステレオ録音された音を制御点において 2 個以上のスピーカーで再生するものである．(図 3.3 参照) この制御は MINT(Multi-input/output INverse Theorem) 法 (多入出力形逆フィルタ) によって制御の有効性が数学的に証明されている．このことにより，家庭の部屋等の再生音場における反響が無視できない場合でも $M + 1$ 個以上のスピーカーを用いることで， M 個の点の音圧を自由に制御することが可能となっている．図 3.4 では制御点 2 個に対してスピーカーを 3 個の場合の構成図である．

3.2 多チャネル-多点制御系

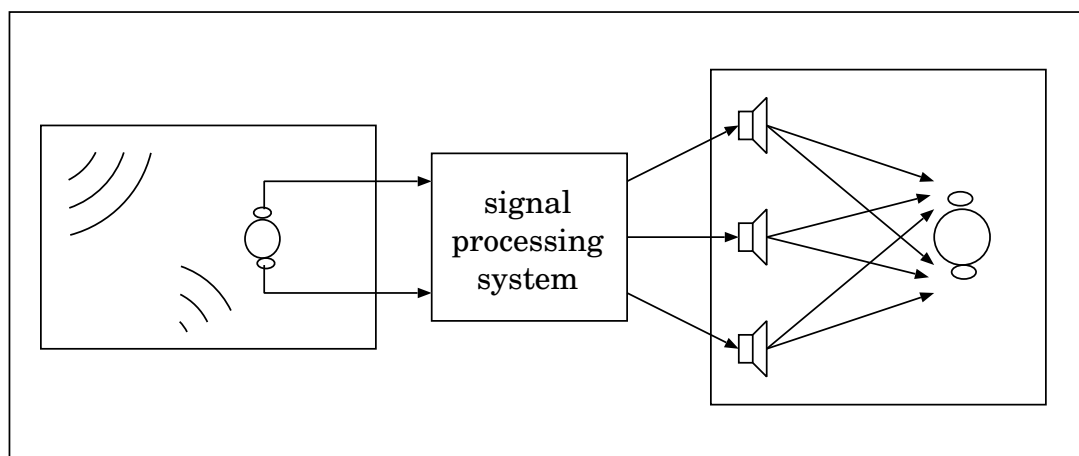


図 3.3 多チャネル-多点制御

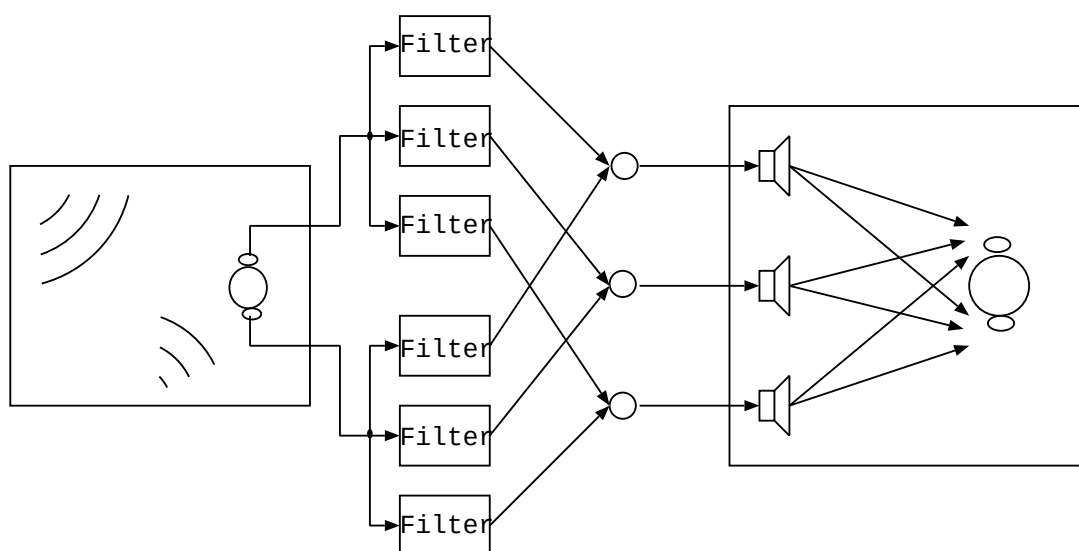


図 3.4 MINT の考え方に基づいた多チャネル-多点制御の構成

3.3 2チャンネル-2点制御系

多チャンネル-多点制御系の問題であるスピーカーが増えることによる制御系の複雑化から計算量の増大があった．それを回避するために浜崎によって提案された2チャンネル-2点制御系について述べる．

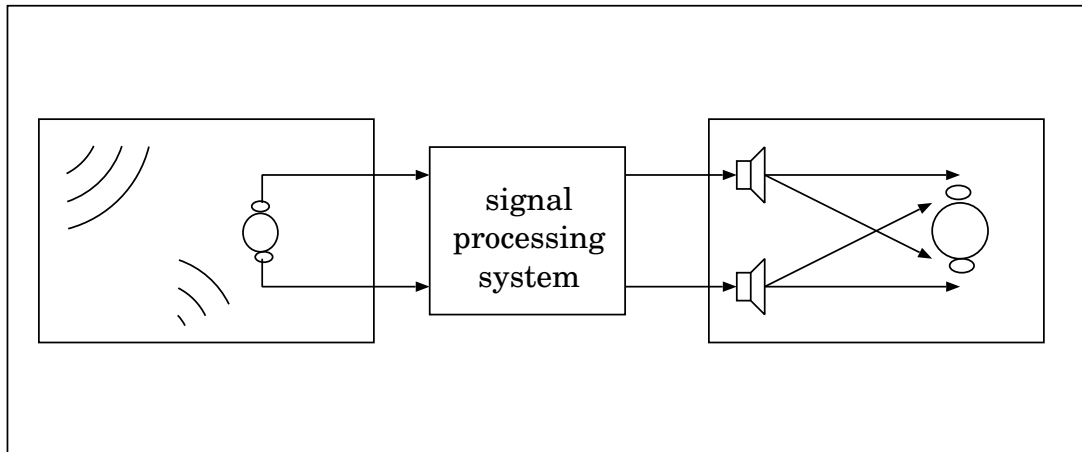


図 3.5 2チャンネル-2点制御

2チャンネル-2点制御系は，ステレオ再生における制御点2個に対してスピーカー2個で行うものである．伝達特性を補正するフィルタが1つの受聴点に入ってくる2本の伝達路に対して1つで済むようになったため，多チャンネル多点制御と比べて少なくなることから問題となっていたスピーカーの増加による演算量の増大を回避している．構成は図3.5のようにシンプルになった．これにより多チャンネル-多点制御と比べてスピーカーの数が減少することで制御系が簡単になり演算量の軽減ができるようになった．この制御系の観測信号は以下のような式で表される．

$$y_1 = x_1 g_{11} + x_2 g_{21} \quad (3.15)$$

$$y_2 = x_2 g_{12} + x_1 g_{22} \quad (3.16)$$

ここで， x_i は入力信号， g_{ij} は室内伝達関数を表す．右側観測信号 y_1 は右側入力信号 x_1 に伝達関数が付加したものと左側入力信号 x_2 に伝達関数が付加したクロストークと呼ばれるものとの総和によって求められる．逆側についても同じことが言える．ここで仮に右側観

3.4 それぞれの問題点

測信号に付加した伝達関数の逆特性が求められたとしてもクロストーク成分があるため一意に観測信号と所望信号が同じになることが言えないことからフィルタの収束特性は未保証である。

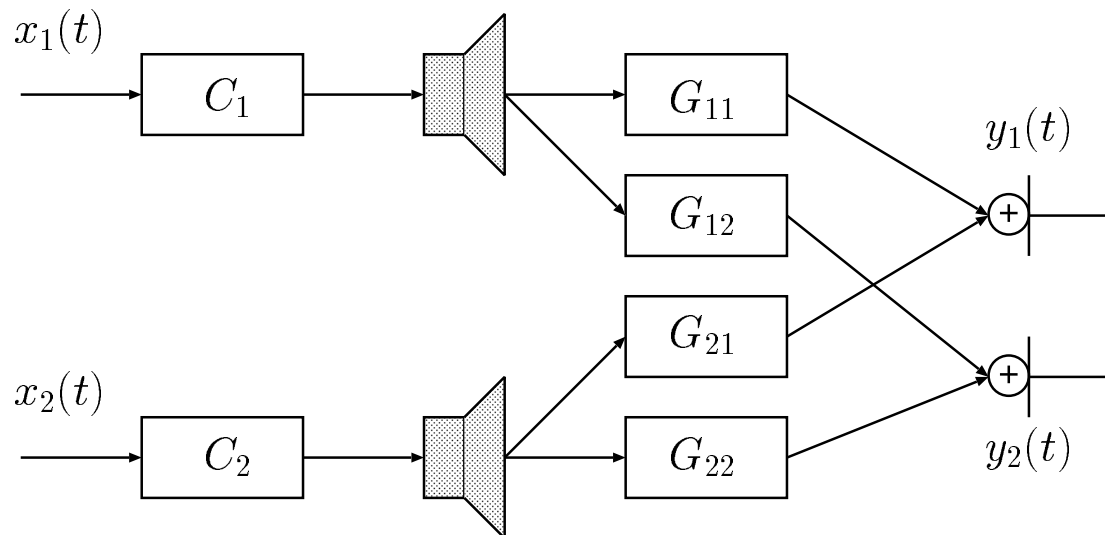


図 3.6 2 チャンネル-2 点制御系のブロック図

3.4 それぞれの問題点

多チャンネル-多点制御の方法として、一般の家庭の環境においても、MINT 法によって M 個の点の音圧を自由に制御できることを先に述べた。しかし、個の方法で逆フィルタを計算し、MINT 法を忠実に再現することは、演算量やハードウェア規模のから考えてこんな場合が多いため現在実用化には至っていない。

2 チャンネル-2 点制御では、演算量とハードウェア規模の問題は解決されているが、音の伝達する過程でクロストークの影響により、観測信号を所望信号に近似させることが難しいという問題がある。

3.5 まとめ

3.5 まとめ

本章ではスピーカーを用いた音場再生でクロストークの問題と補正フィルタの必要性について述べた。また、多チャネル-多点制御系とその問題点を回避するために提案された2チャネル-2点制御系についてとそれらの問題点について述べた。

第 4 章

多入力信号補正システム

4.1 まえがき

2チャンネル-2点制御系の問題点として所望信号と観測信号を十分に近似させるためにはクロストークの影響を軽減する必要があった。そこで、クロストークの問題を解決するために2チャンネル-2点制御系をもとに浜崎によって多入力信号補正システムが提案された。本章では多入力信号補正システムについて説明する。

4.2 多入力信号補正システム

2チャンネル-2点制御系において音が伝達する過程では伝達特性(クロストーク)の影響により受聴点で所望信号を再現するのは難しい。そこで、多入力信号補正システムでは、適応フィルタを用いて近似的に逆特性を推定することで伝達特性の影響を軽減することでクロストークの影響を軽減を行っている。

図 4.1 に多入力信号補正システムの構成を示す。適応フィルタ H_{11} は右側入力信号 $d_1(t)$ とフィルタ通過後の信号 s_{11} から出力信号を算出するようになっている。これより g_{11} の逆特性に近似させるフィルタとなっている。ここで得られた係数を右側補正フィルタ係数 c_1 として与えることによって右側所望信号 x_1 と右側観測信号 y_1 を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り左側観測信号 y_2 を右側所望信号 x_1 に近似させるように作用するため、これを防ぐ必要がある。

適応フィルタ H_{21} は左側観測信号 $d_2(t)$ と $s_{21}(t)$ から出力誤差を算出しているため $g_{21,N}$

4.2 多入力信号補正システム

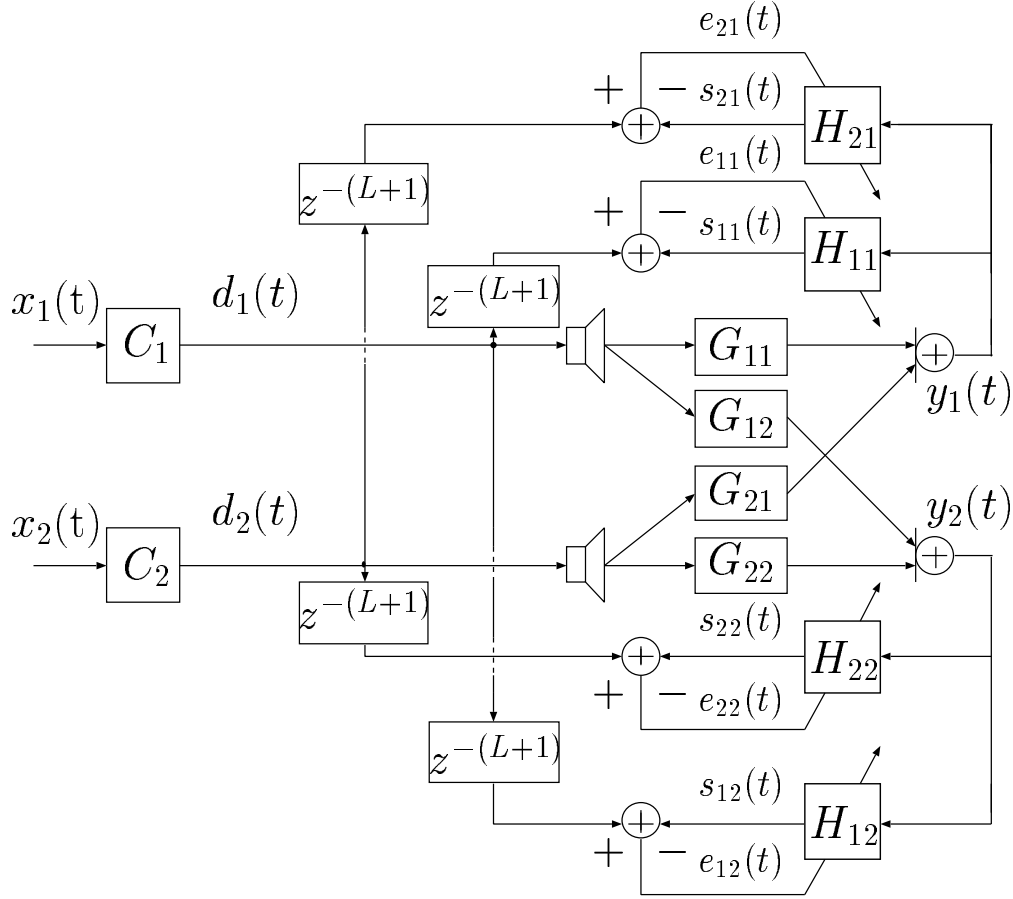


図 4.1 多入力信号補正システム

の逆特性を近似させるフィルタとなる．通常ここで得られるフィルタ係数は左側入力信号から求められたものなので右側補正フィルタ係数 c_1 として与えることはできない．そこで， $g_{12}(k)$ ， $g_{21}(k)$ が互いに強い相関を持つことに着目すると，左側所望信号 $x_2(t)$ から右側観測信号 $y_1(t)$ に到達する信号から求められるフィルタ係数 $\mathbf{h}_{21}(t)$ は右側所望信号 $x_1(t)$ から左側観測信号 $y_2(t)$ に到達する信号の伝達特性を予測するフィルタ $\mathbf{h}_{12}(t)$ として代用できる．これを右側補正フィルタ係数 c_1 として与えることで， $g_{12,N}$ を通過するクロストーク成分を左側所望信号に $x_2(t)$ に近似させることができる．したがって H_{11} で得られたフィルタ係数と H_{21} 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設計できる．

左右の入力信号に対する補正フィルタ係数 $\mathbf{c}_{j,L}$ は

4.3 まとめ

$$\mathbf{c}_{1,L}(t+1) = w_1 \mathbf{h}_{11,L}(t) + (1 - w_1) \mathbf{h}_{21,L}(t) \quad (4.1)$$

$$\mathbf{c}_{2,L}(t+1) = w_2 \mathbf{h}_{22,L}(t) + (1 - w_2) \mathbf{h}_{12,L}(t) \quad (4.2)$$

として与えられる． w_i は第 1 項と第 2 項に対する重み付けパラメータ ($0 \leq w_i \leq 1$) である．
適応フィルタ H_{ij} のフィルタ係数の更新には学習同定法を用い，更新式は，

$$h_{ij,L}(t+1) = h_{ij,L}(t) + \alpha \frac{y_{j,L}(t)}{y_{j,L}(t)^2} e_{ij}(t) \quad (4.3)$$

で示される．ここで， α はステップゲイン， e_{ij} はフィルタの出力誤差を示している．

4.3 まとめ

多入力信号補正システムについて説明した．このシステムでクロストークの影響を軽減するための補正フィルタの精度を決めるパラメータはステップゲイン，重み付けパラメータ等により決定している．

第 5 章

計算機シミュレーション

5.1 まえがき

多入力信号補正システムは観測信号を求める際に室内伝達特性とクロストーク成分における逆特性の問題から一意に所望信号と観測信号を等価できないことから、数学的には解けないため、収束性が未保証である。また、補正フィルタの更新に必要な重み付けパラメータの最適な値も解っていない。そこで、多入力信号補正システムを用いて様々な実環境を想定し、それぞれの環境における伝達関数の実測値を用いて計算機シミュレーションを行うことで、収束特性の検証を行う。

5.2 シミュレーション条件

入力信号として、2 個のマイクロホンでステレオ録音し 8KHz でサンプリングされた成人男性の音声を与える。適応フィルタおよび補正フィルタのインパルス応答長 L を 512、再生空間における伝達関数のインパルス応答長 N を 512、補正フィルタのパラメータ更新に必要なステップゲイン α を 0.05 とした。また、シミュレーション環境として反響性のある部屋においてスピーカ、マイクロホンの位置関係を変えたものを実環境を想定して 7 種類用意した。室内インパルス応答長の測定環境は図 5.2～図 5.8 に示す。

原音に対する再現音の再現精度を、次ぎのような SNR の形で表現する。

$$SNR[dB] = 10 \log_{10} \frac{E[x_i^2(t)]}{E[y_j^2(t) - x_i^2(t)]} \quad (5.1)$$

5.2 シミュレーション条件

ただし, $E[\cdot]$ は期待値, $x_i(t)$ は所望信号の音圧, $y_j(t)$ は観測信号の音圧を表す. また, 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 と多入力信号補正システムを用いて算出した SNR との差を改善量として評価を行う.

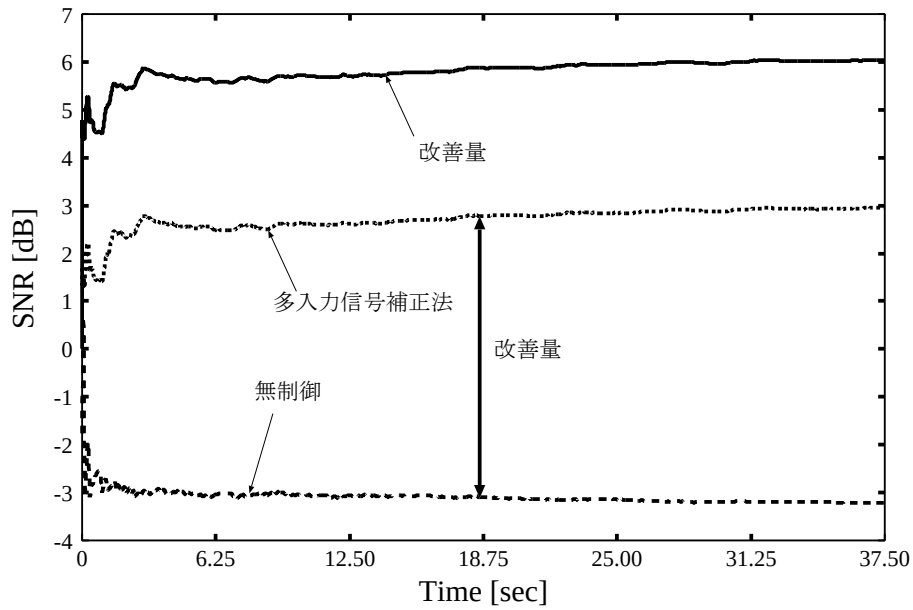


図 5.1 改善量の算出

また, クロストーク成分の相互相関係数 ϕ も評価量とする. ϕ は

$$\phi = \frac{E[g_{21}(k)g_{12}(k)]}{\sqrt{E[g_{21}^2(k)]}\sqrt{E[g_{12}^2(k)]}} \quad (5.2)$$

という式で表され, クロストーク成分である g_{12} と g_{21} がどれだけ近似しているかを調べるものである. この値が高い程クロストーク成分同士が近似していること示している. 言い換えれば, クロストークの伝達特性が強い相関を持つといえる.

補正フィルタ係数を決定する際に, クロストーク同士が強い相関を持つすれば, 一方のクロストーク成分の予測を逆チャネルのクロストークに対応するフィルタで代用するため, よりシステムの精度が上がると思われる.

5.3 シミュレーション結果

5.2 の条件のもとで、重み付けパラメータ w_1, w_2 をそれぞれ 0.05 ずつ 0.00 から 1.00 まで変化させたものの SNR の値から改善量を出し、その中から改善量の値が最大となるもので大きい値から順に並べたものである。それと共にパターンごとの設定環境におけるクロストーク成分の相互相関係数 ϕ も示している。ここで表しているチャンネル 1 は x_1 (図 4.1 における上の信号)、チャンネル 2 は x_2 (下の信号) のことを示している。

表 5.1, 表 5.2 より、クロストーク成分の相互相関係数が大きい値を示す程改善量が大きい値示していることが確認できる。重み付けパラメータからはチャンネル 1 においては w_1 が小さいことから逆チャンネル (チャンネル 2) の所望信号に近似させるように働いていることがわかる。 w_2 が小さいことからクロストーク成分を抑えようとしていることがわかる。次に、チャンネル 2 の重み付けパラメータからは w_1 が小さいことからクロストーク成分を抑えようとしていることがわかる。 w_2 が大きいことからチャンネル本来の信号を大きく影響させようとしていることがわかる。

これは、チャンネル 2 においては w_1 からはクロストーク成分を抑制し、 w_2 からはチャンネル本来の信号に影響させようとしているため、左側観測信号と左側所望信号を近似させようとしているためだと考えられる。しかし、チャンネル 1 においては w_2 からクロストーク成分を抑制しようとしていることで右側観測信号を右側所望信号に近似させようとしている傾向が伺えるが、 w_1 では本来ではチャンネル本来の信号に影響させることが望ましいがそのようには影響していない。

この原因として考えられるのは、所望信号のパワー比を取ったときに、チャンネル 2 よりチャンネル 1 のほうが遥かに大きかったことから、それほど補正をしなくてもチャンネル 1 の音声信号は所望信号と観測信号が近似しているのではないかと考えられる。

このことから、重み付けパラメータの設定には所望信号のパワー比に関係があることが確認できる。

5.4 まとめ

本章では，実環境を想定した様々なパターンに対してシミュレーションによって求めた伝達関数を用いてクロストーク成分における相互相関係数と重み付けパラメータの変化から多入力信号補正システムにおける収束特性を検証した．

結果として，クロストーク成分における相互相関係数が高い程改善量が高くなりシステムが良好な結果を示すことが確認できた．また，入力信号のパワー比による重み付けパラメータの収束特性が見られた．

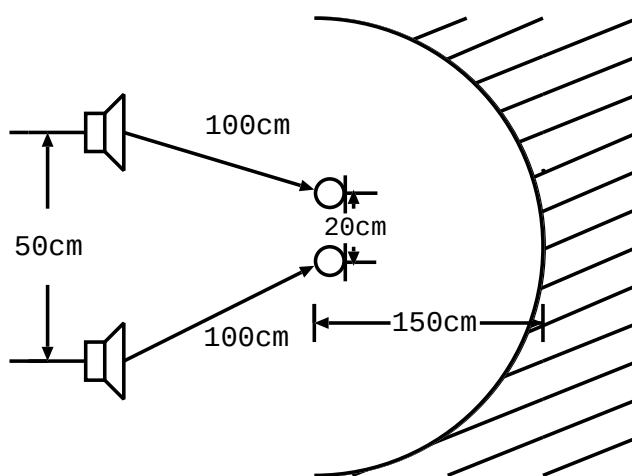


図 5.2 パターン 1

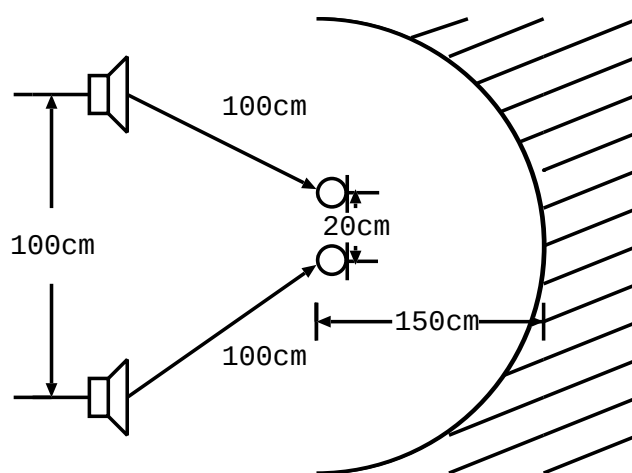


図 5.3 パターン 2

5.4 まとめ

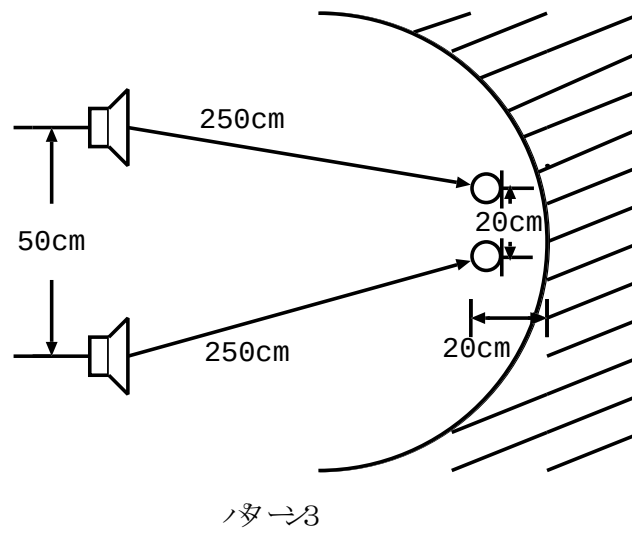


図 5.4 パターン 3

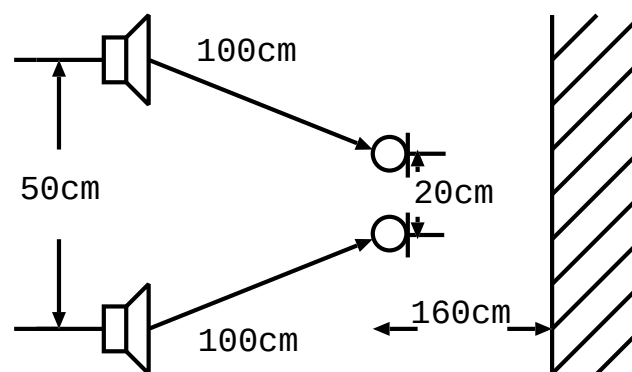


図 5.5 パターン 4

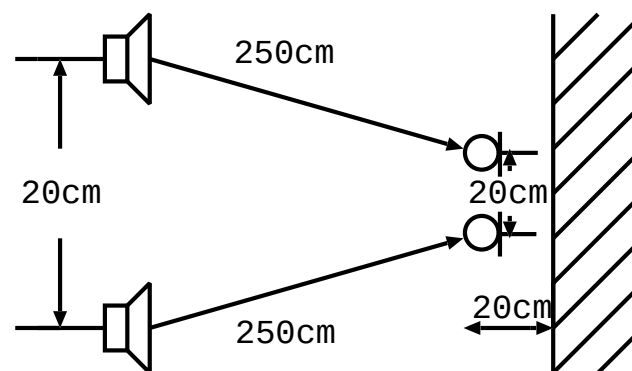


図 5.6 パターン 5

5.4 まとめ

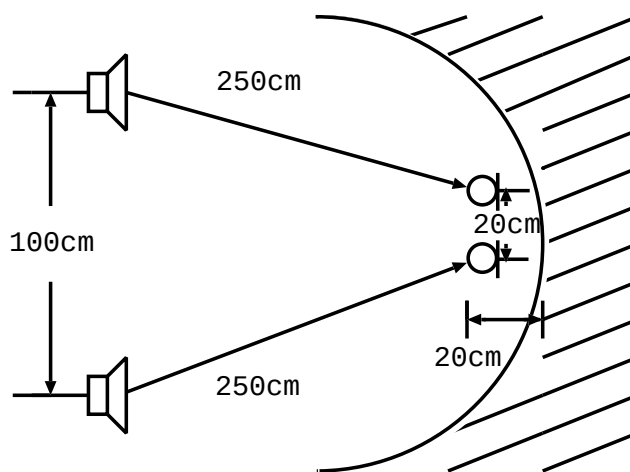


図 5.7 パターン 6

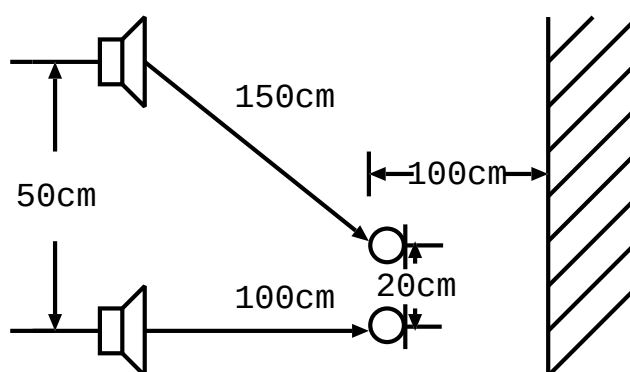


図 5.8 パターン 7

5.4 まとめ

表 5.1 チャンネル 1 のシミュレーション結果

	ϕ	w_1	w_2	改善量[dB]
パターン 1	0.90	0.25	0.20	5.99
パターン 2	0.87	0.25	0.35	5.73
パターン 3	0.82	0.25	0.20	5.64
パターン 4	0.80	0.15	0.35	5.51
パターン 5	0.76	0.15	0.20	5.29
パターン 6	0.70	0.25	0.50	5.20
パターン 7	0.56	0.25	0.35	4.93

表 5.2 チャンネル 2 のシミュレーション結果

	ϕ	w_1	w_2	改善量[dB]
パターン 1	0.90	0.15	0.80	5.96
パターン 2	0.87	0.25	0.75	5.93
パターン 3	0.82	0.20	0.80	5.65
パターン 4	0.80	0.20	0.80	5.42
パターン 5	0.76	0.15	0.80	5.31
パターン 6	0.70	0.20	0.80	5.27
パターン 7	0.56	0.25	0.75	4.96

第 6 章

結論

6.1 本研究における評価

原音場の特性を再生音場にて忠実に再現する手法として多入力信号補正システムが提案されたが，伝達特性の近似逆特性の収束性について保証がされていなかった．そこで，実環境を想定し様々なパターンの室内インパルス応答を用いて重み付けパラメータとクロストーク成分の相互相関から多入力信号補正システムの収束特性を検証した．

計算機シミュレーションにより，クロストーク成分における相互相関係数が高い程改善量が高くなることがわかり，システムの有効性が検証された．また，入力信号のパワー比によって重み付けパラメータの収束特性が見られた．

6.2 今後の課題

本論文では多入力信号補正システムにおける重み付けパラメータのパワー比による傾向を示した．しかし，まだパラメータにはばらつきがあり最適な値として与えることはできない．このことから，最適な重み付けパラメータの値を検証するには入力信号のパワー比を同じにすることで伝達特性のパワー比から検討する必要がある．

謝辞

本研究において，御指導頂いた福本助教授に感謝致します．

参考文献

- [1] 秋山由佳, 福本昌弘, “聴覚特性にもとづいた重みつき誤差による多入力信号補正システム,” 高知工科大卒業論文, 2002.
- [2] 浜崎真二, 福本昌弘, “クロストーク成分における相互相関係数に着目した音場再生システム,” 信学技報, SIP2004-115, pp.31–36, 2005 年 1 月.
- [3] 大賀寿郎, 山崎芳男, 金田豊, “音響システムとデジタル信号処理,” 電子情報通信学会, 1995.
- [4] 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.
- [5] S. ヘイキン (武部幹 訳), “適応フィルタ入門,” 現代工学社, 1994.
- [6] D.H.Cooper and J.L.Bauck, “Prospect for transaural recording,” J.Audio Eng. Soc. ,vol.37, 1/2.