

平成 21 年度
プロジェクト研究報告書

信号補正システムによる音場再生の 性能評価

1100305 四宮 隼人

指導教員 福本 昌弘

2010 年 3 月 1 日

高知工科大学 情報システム工学科

要 旨

信号補正システムによる音場再生の性能評価

四宮 隼人

原音場の特性を異なる空間で再現するためには，伝達特性の影響を考慮しなければならない．スピーカを用いた音場再生システムであるトランスオーラルシステムは受音点で伝達特性の影響を受ける．一般に1つのスピーカから発せられた音は左右の耳に2つの異なった伝達経路をとる．1つのスピーカでは1つの補正フィルタに対する処理しか行うことができず，複数の伝達経路からくるそれぞれの伝達特性を補正するにはその数だけスピーカが必要になり制御系が複雑になる．この問題に対して，1つの補正フィルタで2つの伝達経路を制御する多入力信号補正システムが浜崎らによって提案された．このシステムはクロストーク成分が対称である理想的な環境の下で有効性が確認されている．しかし，受聴者が移動した時等，理想的ではない環境でのシステムの有効性は保証されていない．

本研究では，左右のクロストーク成分の異なる環境における伝達関数を想定し，左右のクロストーク成分の違いが大きい程原音に対する再現音の精度が減少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 適応フィルタ，音場再生，多入力信号補正システム，クロストーク

目次

第 1 章	序論	1
1.1	本研究の背景と目的	1
1.2	本論文の概要	2
第 2 章	音響モデルにおけるデジタル信号処理	3
2.1	まえがき	3
2.2	FIR デジタルフィルタ	4
2.3	適応信号処理	5
2.3.1	適応フィルタ	5
2.3.2	適応アルゴリズム	7
2.3.3	学習同定法	8
2.4	まとめ	9
第 3 章	音場再生システム	10
3.1	まえがき	10
3.2	バイノーラルシステム	10
3.3	トランスオーラルシステム	11
3.4	多チャネル-多点制御系	13
3.5	2 チャネル-2 点制御系	14
3.6	問題点	15
3.7	まとめ	16
第 4 章	多入力信号補正システム	17
4.1	まえがき	17
4.2	多入力信号補正システム	17

目次

4.3	まとめ	19
第 5 章	計算機シミュレーション	20
5.1	まえがき	20
5.2	シミュレーション条件	20
5.3	シミュレーション結果	21
5.4	まとめ	22
第 6 章	結論	28
6.1	本研究における評価	28
6.2	今後の課題	28
	謝辞	29
	参考文献	30

目次

2.1	FIR デジタルフィルタ	4
2.2	適応フィルタを用いたブロック図	6
3.1	室内伝達特性の補正	11
3.2	トランスオーラルシステムの模式図	12
3.3	MINT 法に基づく多チャネル-多点制御の構成	14
3.4	2 チャネル-2 点制御	14
4.1	多入力信号補正システム	18
5.1	左側音声信号	23
5.2	右側音声信号	23
5.3	左側のインパルス応答	24
5.4	右側のインパルス応答	24
5.5	左側のインパルス応答	25
5.6	右側のインパルス応答	25
5.7	左側のインパルス応答	26
5.8	右側のインパルス応答	26
5.9	それぞれの環境での改善量の収束特性	27

表目次

第 1 章

序論

1.1 本研究の背景と目的

近年，デジタル技術の著しい発展から，マルチメディアの時代と言われはじめて久しい．音声や映像など様々なデータの伝送が簡単に可能になり，デジタル信号処理が現代の情報・通信産業において重要な役割をなしている．

デジタル信号処理の中でも，バーチャルリアリティを実現する試みが各方面で行われている．バーチャルリアリティの実現として音響の分野でも元の空間での音場を再現空間で実現できる音場再生システムの構築が求められている．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，左右の耳に対して2つの異なった伝達経路をとる．それぞれの伝達特性の影響を除去するためには2つの補正フィルタが必要となる．しかし，1つのスピーカでは1つの補正フィルタに対する処理しかできないため，複数経路に対する処理を単一制御系で行う必要が生じる．

この問題に対し，浜崎らによって1つのフィルタにより2つの伝達経路を制御する多入力信号補正システムが提案された．このシステムは受聴者の移動によって生じるクロストーク部分の伝達特性の変動を考慮しない場合に有効性が確認されている．しかし，受聴者が移動した時等，伝達特性の変動が生じた場合のシステムの有効性は保証されていない．

本研究では，クロストーク成分における近似性の異なる環境の伝達関数を想定し，計算機シミュレーションを行うことによってクロストーク成分の伝達関数に違いが生じると改善量にどのような影響を与えるのか検証を行う．

1.2 本論文の概要

本論文の概要について述べる．

第2章では，デジタル信号処理の基礎技術である FIR デジタルフィルタと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，今回評価の対象と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，音場再生技術であるバイノーラルシステムとトランスオーラルシステムについて述べる．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ついて述べる．

第5章では，4章で述べたシステムの様々な環境を想定した計算機シミュレーションを行うことによって，それぞれの環境での改善量の検証を行う．

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，今後の課題についての考えを述べる．

第 2 章

音響モデルにおけるデジタル信号処理

2.1 まえがき

デジタル信号処理で扱うものとして、音楽、音声などがある。これら音響信号は連続的な量をも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も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。

様々な信号処理アルゴリズムの中で、デジタルフィルタが最もよく利用されている。デジタルフィルタとは複雑な信号波形から所望の信号を抽出するために用いられる。処理の対象となる観測信号は、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、観測信号から目的とする信号成分を抽出するためにデジタルフィルタを用いる。

デジタルフィルタは、フィードバックのループがなく、有限長の応答をもつ FIR(Finite Impulse Response) フィルタと、フィードバックのループがあり、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、評価対象とする多入力信号補正システムが有限長のフィルタを用いていることから、FIR デジタルフィルタについて述べる。そして FIR デジタルフィルタを用いて未知のパラメータを推定する機能をもつ適応フィルタについての説明を行う。

2.2 FIR デジタルフィルタ

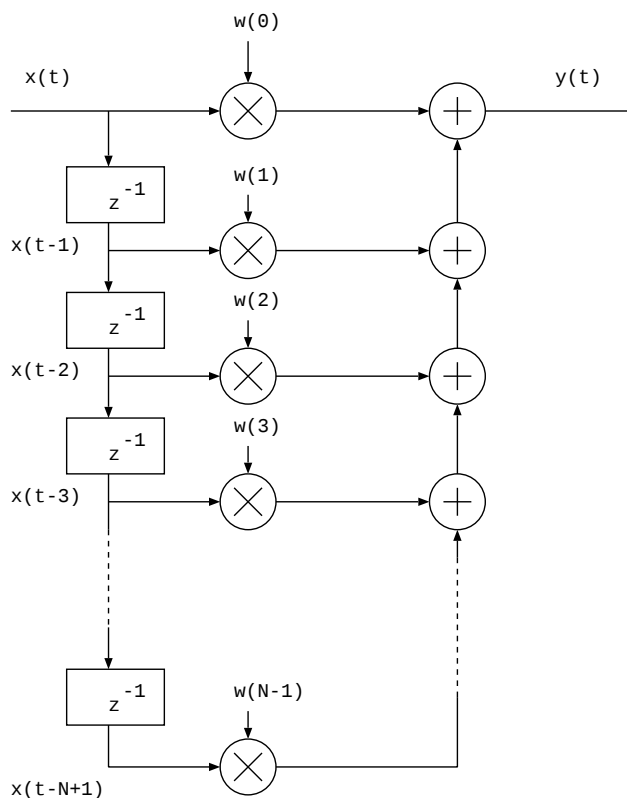


図 2.1 FIR デジタルフィルタ

2.2 FIR デジタルフィルタ

FIR(Finite Impulse Response) デジタルフィルタとはインパルスを入力した時の出力信号が有限時間で 0 に収束するフィルタのことである．図 2.1 は，単位遅延素子 Z^{-1} を複数用いて標本化周期ごとに入力された N 個の信号を加算して，出力する構成のフィルタである．入力信号 $x(t)$ は単位遅延を受けるごとに乗算器をもつ各タップに出力され，パラメータ (フィルタ係数) w_N が乗算される．その後それらが加算された結果 $y(t)$ が出力される．

このフィルタの入出力関係は，

$$y(t) = \sum_{i=0}^{N-1} w(i)x(t-i+1) \quad (2.1)$$

で与えられる．

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

2.3 適応信号処理

式 (2.1) から、出力信号 $y(t)$ はパラメータ w_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れば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 は、

$$w_N = [w(0), w(1), w(2), \dots, w(N)]^T \quad (2.2)$$

で表される。ここでの T は転置を表す。

FIR フィルタは IIR フィルタに比べて演算誤差の影響が小さく、また有限の振幅の入力信号に対して出力信号の振幅も有限になり、理論的には必ず安定になるので比較的扱いやすいフィルタである [3]。

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、時間と共に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。このような処理を行うために、Widrow らの先駆的な研究により始められたのが適応信号処理である [8]。

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。本論文で、 t は時刻、 N はインパルス応答長を示す。ここで、 $x(t)$ は入力信号、 $d(t)$ は未知系の出力信号、 $y(t)$ は推定系の出力信号、 $e(t)$ は $d(t)$ と $y(t)$ の出力誤差を表す。また、 w_N は未知システムのフィルタ係数、 h_N は適応フィルタのフィルタ係数を表す。パラメータ h_N は、

$$h_N = [h(0), h(1), h(2), \dots, h(N)]^T \quad (2.3)$$

2.3 適応信号処理

で表される．

このとき，未知系出力信号 $d(t)$ と推定系出力信号 $y(t)$ は

$$d(t) = \sum_{i=0}^{N-1} w(i)x(t-i+1) \quad (2.4)$$

$$y(t) = \sum_{i=0}^{N-1} h(i)x(t-i+1) \quad (2.5)$$

と表すことができる．また，出力誤差 $e(t)$ は

$$e(t) = d(t) - y(t) \quad (2.6)$$

と表される．

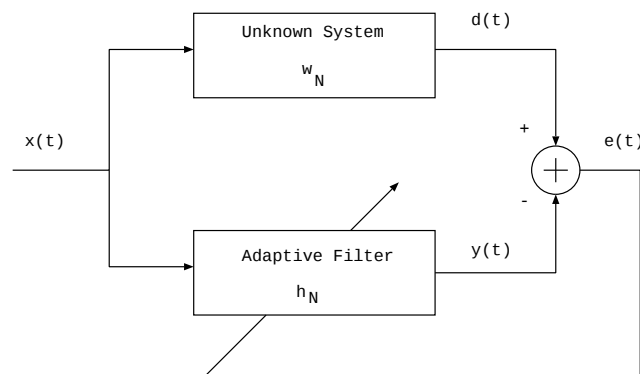


図 2.2 適応フィルタを用いたブロック図

式 (2.4)，(2.5) において

$$w(i) = h(i) \quad ; \forall i = 0, 1, 2, \dots, N-1 \quad (2.7)$$

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．

しかし，一般に未知システムのインパルス応答長は無限に続く．この場合，式 (2.7) のような仮定が成り立たないために，同一の入力信号に対して完全に等しい出力結果が得られるような FIR デジタルフィルタを構成することはできない．

2.3 適応信号処理

そこで、室内音響系をはじめとする実際のシステムの多くのインパルス応答が、時間と共に減衰していくという性質を利用する。

すると、未知系システムのインパルス応答のうち最初の適当な N 個の値を推定することによって、推定系出力信号を未知系出力信号に近づけることが可能である。

このようにして、推定システムの出力 $y(t)$ が未知システムの出力 $d(t)$ に近似するように推定システムのパラメータ h_N を逐次的に推定する学習機能をもったフィルタのことを、適応フィルタという [4, 6, 8]。

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 $x(t)$ 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段のことである [6, 8]。

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 $x(t)$ に対する推定系出力 $y(t)$ は

$$y(t) = \sum_{i=0}^{N-1} h(i)x(t-i+1) \quad (2.8)$$

のように与えられた。式 (2.8) で表される $y(t)$ は、入力状態ベクトル $x_N(t)$ および係数ベクトル h_N の内積として

$$y(t) = h_N^T x_N(t) \quad (2.9)$$

と表すことができる。ただし、 x_N , h_N はそれぞれ

$$h_N = [h(0), h(1), h(2), \dots, h(N-1)]^T \quad (2.10)$$

$$x_N = [x(1), x(2), x(3), \dots, x(N-1)]^T \quad (2.11)$$

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

$$J = E[e^2(t)]$$

2.3 適応信号処理

$$\begin{aligned} &= E[(d(t) - y(t))^2] \\ &= E[(d(t) - h_N^T x_N(t))^2] \end{aligned} \quad (2.12)$$

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 [8] 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ため演算量は LMS アルゴリズムより多少増加する．しかし，高速な収束特性を有しており，入力信号の大きさによって，収束特性が変化するという問題点がなく，実用的にも優れた適応アルゴリズムである．学習同定法は別名 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 t において，推定系出力 $y(t)$ が未知系出力 $d(t)$ に等しいとすると

$$d(t) = h_N^T x_N(t) \quad (2.13)$$

と表すことができる．

しかし， $h_N = w_N$ を満たすためには，すべての入力信号 $x(t)$ に対して式 (2.13) が成り立たなければならない．

そこで，式 (2.13) 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 (2.13) より，入力ベクトル $x_N(t)$ に直交しているといえる．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 $x_N(t)$ 方向にパラメータ修正したとき，もっとも w_N に近い点といえる．

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

$$\begin{aligned} h_N(t+1) &= h_N(t) + (h_N(t+1) - h_N(t)) \\ &= h_N(t) + \frac{(w_N - h_N(t))^T (h_N(t+1) - h_N(t))}{\|h_N(t+1) - h_N(t)\|} \\ &\quad \cdot \frac{h_N(t+1) - h_N(t)}{\|h_N(t+1) - h_N(t)\|} \end{aligned} \quad (2.14)$$

2.4 まとめ

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 (2.14) において

$$\frac{(w_N - h_N(t))^T (h_N(t+1) - h_N(t))}{\|h_N(t+1) - h_N(t)\|}$$

はパラメータの修正量を示し

$$\frac{h_N(t+1) - h_N(t)}{\|h_N(t+1) - h_N(t)\|}$$

はパラメータの修正方向を示す。

ここで

$$\frac{h_N(t+1) - h_N(t)}{\|h_N(t+1) - h_N(t)\|} = \frac{x_N(t)}{\|x_N(t)\|} \quad (2.15)$$

$$\begin{aligned} (w_N - h_N(t))^T x_N(t) &= d(t) - y(t) \\ &= e(t) \end{aligned} \quad (2.16)$$

が成立するので、式 (2.14) は

$$h_N(t+1) = h_N(t) + \frac{x_N(t)}{\|x_N(t)\|^2} e(t) \quad (2.17)$$

と変形できる。

学習同定法は、式 (2.17) の修正ベクトルにステップゲインを掛け

$$h_N(t+1) = h_N(t) + \alpha \frac{x_N(t)}{\|x_N(t)\|^2} e(t) \quad (2.18)$$

で与えられる。

2.4 まとめ

本章では、FIR デジタルフィルタ、適応フィルタ、適応アルゴリズムなどのデジタル信号処理について述べた。また、多入力信号補正システムに利用されている適応アルゴリズムである学習同定法について述べた。

第 3 章

音場再生システム

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原から直接到達する音の他に，壁などで反射された反響音，残響音により構成されている。この反響，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。

現在，音場再生を実現する方法として，ヘッドフォンを用いて再生するバイノーラルシステム技術とスピーカを用いて再生するトランスオーラルシステム技術が既に研究されている。

本章では，バイノーラルシステムとトランスオーラルシステムについて述べ，スピーカ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。

3.2 バイノーラルシステム

バイノーラルシステムとは，原音場における音場をヘッドフォンを利用することによって再現する技術である [5, 6]。

原音場にダミーヘッドを設置し，ダミーヘッドの左右の耳の位置にマイクロフォンを用い

3.3 トランスオーラルシステム

て2チャンネル録音を行う。収音された音響信号は、記録・伝送などを経た後ヘッドフォンを用いて受聴者の耳元で再生される。よって再生された音響信号は、再現空間における室内の伝達特性の影響を受けることがない。このため再現空間では、あたかも原音場で受聴しているような状況を実現する。

3.3 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、スピーカを用いて音場再生を実現しようとする技術である [5, 6]。原音場にダミーヘッドを設置して2チャンネル収音された音は、再現空間においてスピーカを用いて再生される。スピーカ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。

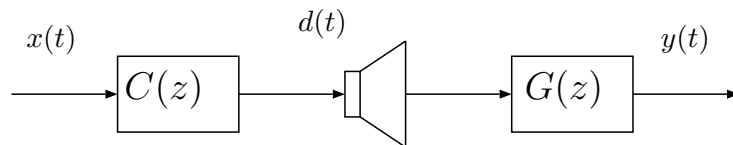


図 3.1 室内伝達特性の補正

室内伝達特性 $G(z)$ を補正するモデル図 3.1 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 $C(z)$ を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

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

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

となる。式 (3.1)(3.2) より

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

3.3 トランスオーラルシステム

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 3.2 に示す。

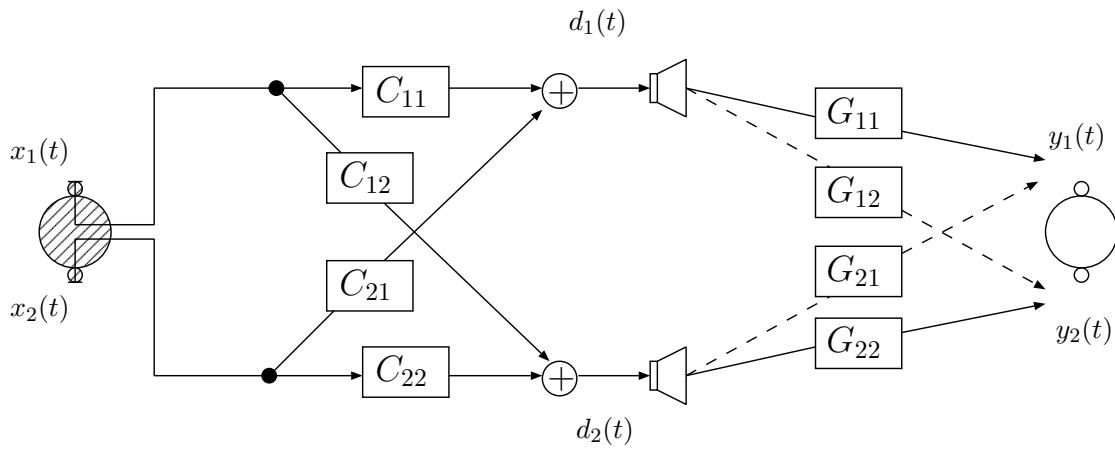


図 3.2 トランスオーラルシステムの模式図

このとき、観測信号と入力信号の関係は

$$Y_1(z) = G_{11}(z)D_1(z) + G_{21}(z)D_2(z) \quad (3.5)$$

$$Y_2(z) = G_{12}(z)D_1(z) + G_{22}(z)D_2(z) \quad (3.6)$$

となり行列

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

で表すことができる。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象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

3.4 多チャネル-多点制御系

を再現するための妨害要因となる．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る．所望信号と入力信号の関係を式 (3.7) と同様に行列で表すと

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

となる．室内伝達経路系，補正フィルタ系をそれぞれ

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

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

とおく．また

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

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

とおくと式 (3.7)(3.8) より

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

となり

$$G(z)C(z) = 1 \quad (3.14)$$

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．

3.4 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 個以上のスピーカで再生するものである．この制御は MINT(Multi-input/output INverse Theorem) 法 (多入出力系逆フィルタ) によって制御の有効性が数学的に証明されている [7]．これによ

3.5 2チャンネル-2点制御系

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 3.3 では制御点 2 個に対してスピーカを 3 個の場合の構成図である．

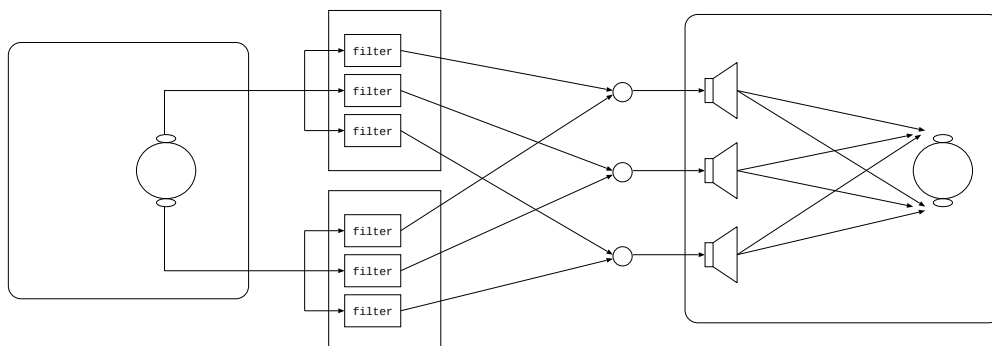


図 3.3 MINT 法に基づく多チャンネル-多点制御の構成

3.5 2チャンネル-2点制御系

多チャンネル-多点制御系の問題として，スピーカ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 [1]．

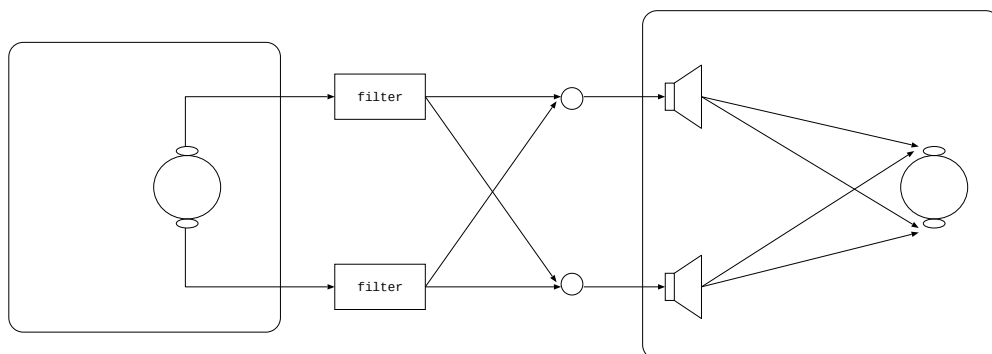


図 3.4 2チャンネル-2点制御

2 チャンネル-2 点制御系は，ステレオ録音された音を制御点 2 個に対してスピーカ 2 個で再生を行うものである．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

3.6 問題点

達路に対して1つで済むようになったため、多チャネル-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の増加による演算量の増大を回避している。構成は図 3.4 のようにシンプルになった。これにより多チャネル-多点制御と比べてスピーカの数が増減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。

3.6 問題点

多チャネル-多点制御と2チャネル-2点制御にはそれぞれ以下の問題点がある。

多チャネル-多点制御では、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述べた。しかし、演算量やハードウェア規模から考えてMINT法を忠実に再現することは困難な場合が多いため現在実用化には至っていない。2チャネル-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。

3.7 まとめ

3.7 まとめ

本章では，スピーカ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，それらの制御系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を行う．

4.2 多入力信号補正システム

2 チャネル-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クロストークの影響の軽減を行っている．

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．これにより 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} 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

4.2 多入力信号補正システム

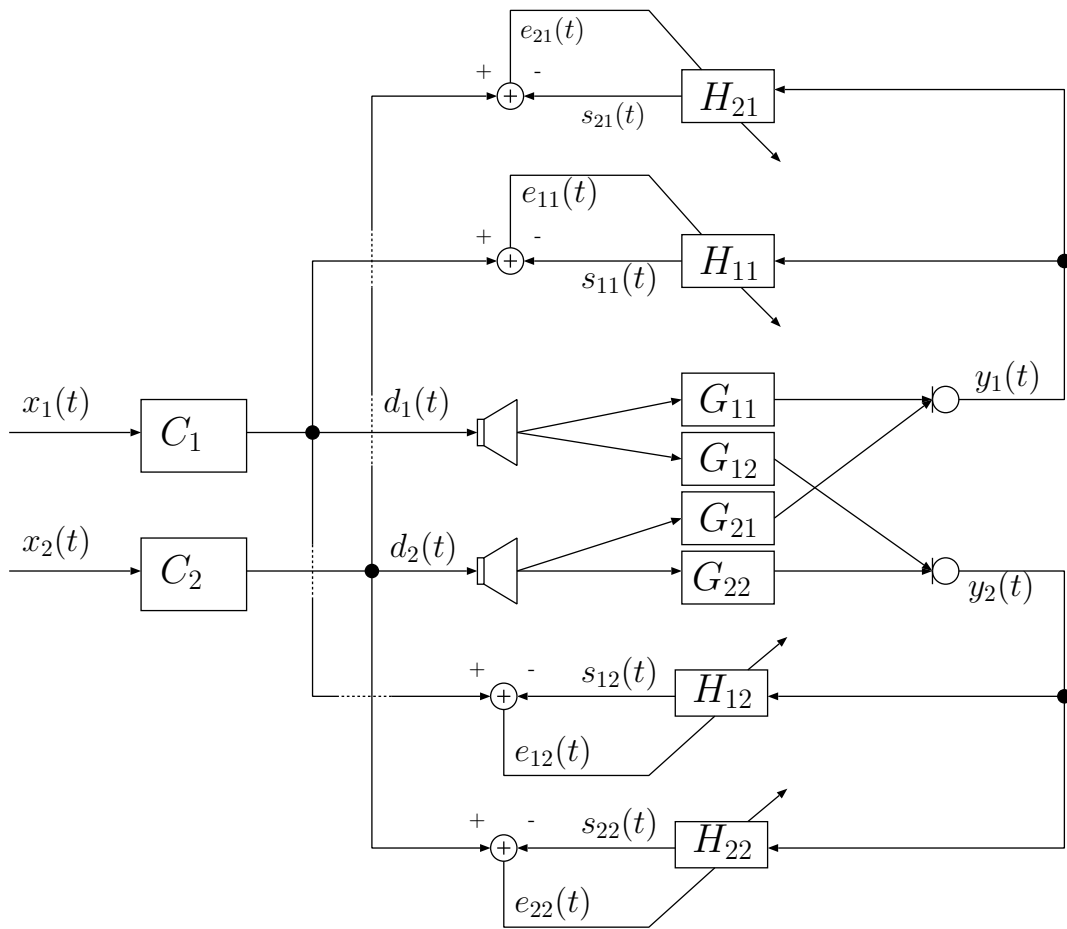


図 4.1 多入力信号補正システム

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} ， g_{21} 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 $h_{21}(t)$ 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 $h_{12}(t)$ 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} 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．したがって適応フィルタ 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持つ補正フィルタが設計できる．

4.3 まとめ

左右の入力信号に対する補正フィルタ係数 c_j は

$$c_1(t+1) = rh_{11}(t) + (1-r)h_{21}(t) \quad (4.1)$$

$$c_2(t+1) = rh_{22}(t) + (1-r)h_{12}(t) \quad (4.2)$$

として与えられる． $r(0 \leq r \leq 1)$ は左右における補正フィルタのパラメータに対する，急激なパラメータ変動を制御する修正量を表す．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

$$h_{ij}(t+1) = h_{ij}(t) + \alpha \frac{y_j(t)}{y_j(t)^2} e_{ij}(t) \quad (4.3)$$

で示される．ここで， α はパラメータ修正の大きさを制御するステップゲイン， e_{ij} は出力誤差を表す．

4.3 まとめ

本章では，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，修正率等によって決定している．

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受聴者が動くこと等による伝達特性の変化がないという理想的な環境と仮定することでクロストーク成分における逆特性の問題を解決している．よって，例えば受聴者が移動した時等，左右のクロストーク成分における伝達特性が変化した場合，適した補正が行われるか未保証である．そこで，左右のクロストーク成分に違いがある環境を想定し，計算機シミュレーションを行うことで左右のクロストーク成分の違いが補正精度にどれほどの影響を与えるか検証を行う．

5.2 シミュレーション条件

入力信号として 44.1kHz でサンプリングされたステレオ音声信号 (図 5.1, 図 5.2) 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を 256，再生空間における伝達関数のインパルス応答長を 128，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を 0.005，修正率を 0.5 とした．またシミュレーション環境としてクロストークにおける伝達関数の左側に変化をつけることで以下の環境を想定した．

1. 理想的な環境

左右のインパルス応答が同じ

2. クロストーク成分が近似している環境

左右のインパルス応答が似ている

5.3 シミュレーション結果

3. クロストーク成分が大きく違う環境

他の想定した環境に比べ、違いが大きい

左右の観測地点へのインパルス応答には、各々独立に発生された乱数が指数関数的に減少するように重みをかけたものを用いた。それぞれ想定した環境のクロストーク成分の室内インパルス応答を図 5.3～図 5.8 に示す。

評価量には ERLE(Echo Return Loss Enhancement) を用いる。原音に対する再現音の再現精度を、次のような式で表現する。

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

ここで、 $E[\cdot]$ は期待値、 $x_i(t)$ は所望信号、 $y_j(t)$ は観測信号を表す。

5.3 シミュレーション結果

今回行った条件のもとで、シミュレーションを行った結果を図 5.9 に示す。図 5.9 はサンプル毎の改善量を示しており、原信号と補正された信号との誤差の減少に伴い改善量が上昇する。この改善量が高い程補正された信号が原信号に近いということになる。

今回シミュレーションした環境の内、理想的な環境では左右のクロストーク成分が対称であるためクロストーク問題は解決しているといえる。この環境はクロストーク問題を解決しているというシステムにとって理想的な環境であるため、補正された信号は最も原信号に近くなる。

クロストーク成分の近似している環境ではクロストーク問題はある程度軽減していることになるので改善量は理想的な環境の改善量より少し低い結果となった。このことから、左右のクロストーク成分が近似している程改善量が高くなっていくと考えられる。

クロストーク成分が大きく違う環境では、本来観測されるべき左右の観測信号はクロストークの影響が大きく改善量は近似していない分、さらに減少している結果となった。このことから、左右のクロストーク成分の違いが大きくなる程改善量が低くなっていくと考えられる。

5.4 まとめ

以上の結果から、クロストーク成分が対称ではない環境では左右のクロストーク成分の違いが大きくなるほど信号補正における改善量が下がっていき、左右のクロストーク成分が近似するほど改善量が上がっていくと考えられる。したがって、理想的な環境と違いが信号補正精度に影響することがわかる。

5.4 まとめ

本章では、クロストーク成分が同じ理想的な環境とクロストーク成分の近似している環境、他に比べクロストーク成分が大きく違う環境を想定した伝達関数を用いて多入力信号補正システムにおける信号補正の改善量を検証した。

結果として、クロストーク成分が対称ではない環境では左右のクロストーク成分の違いが大きくなる程改善量が減少していくことが確認できた。

5.4 まとめ

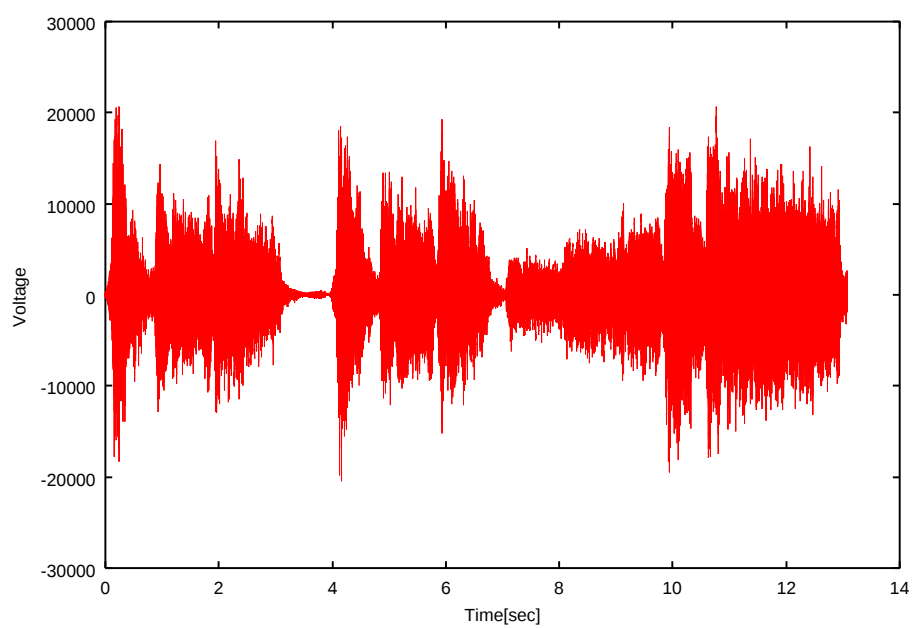


図 5.1 左側音声信号

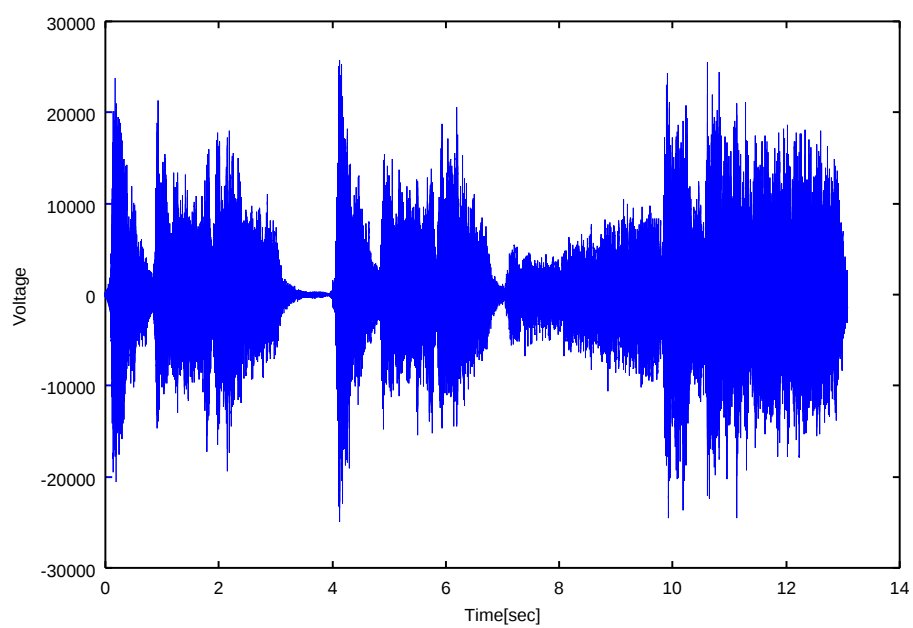


図 5.2 右側音声信号

5.4 まとめ

理想的な環境

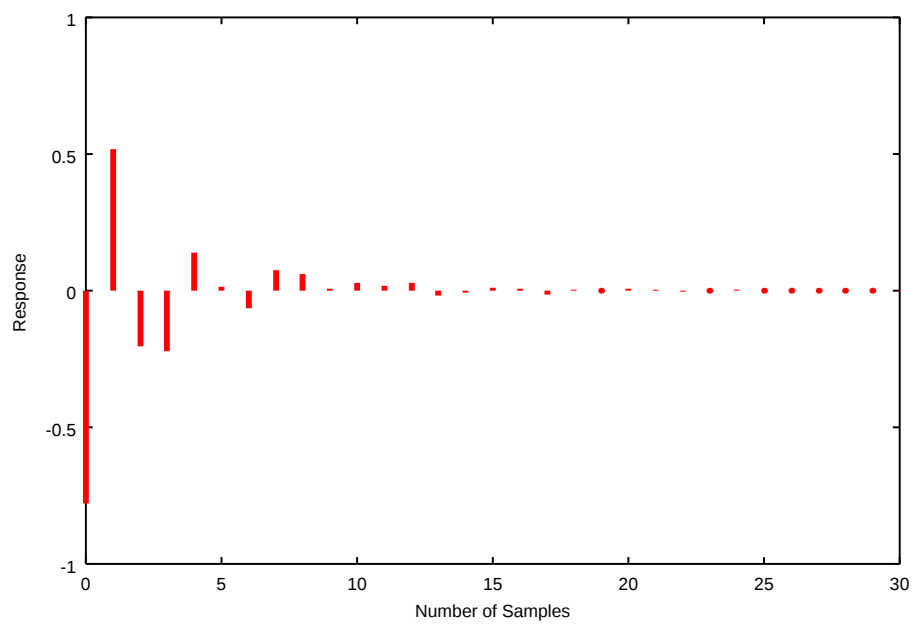


図 5.3 左側のインパルス応答

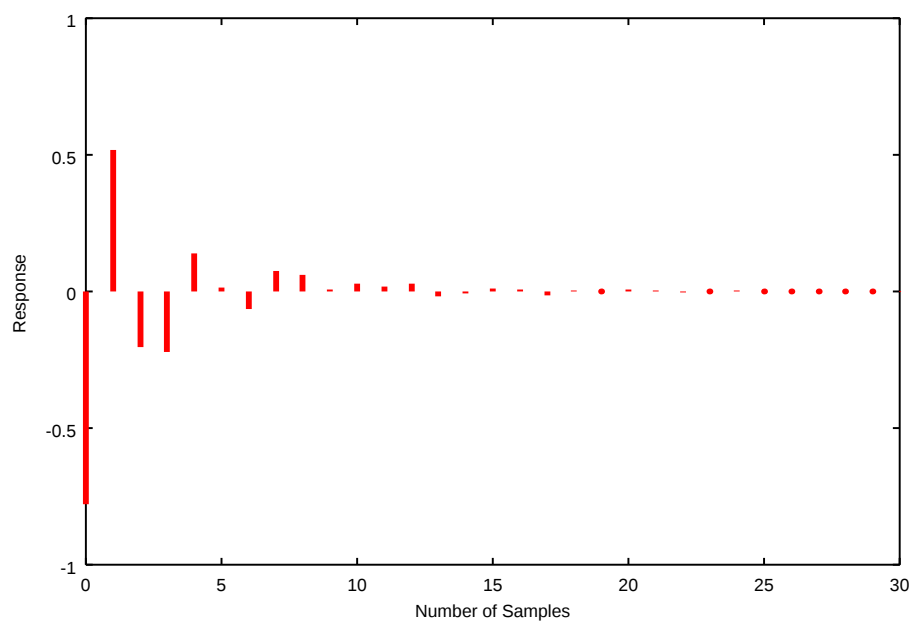


図 5.4 右側のインパルス応答

5.4 まとめ

クロストーク成分が近似している環境

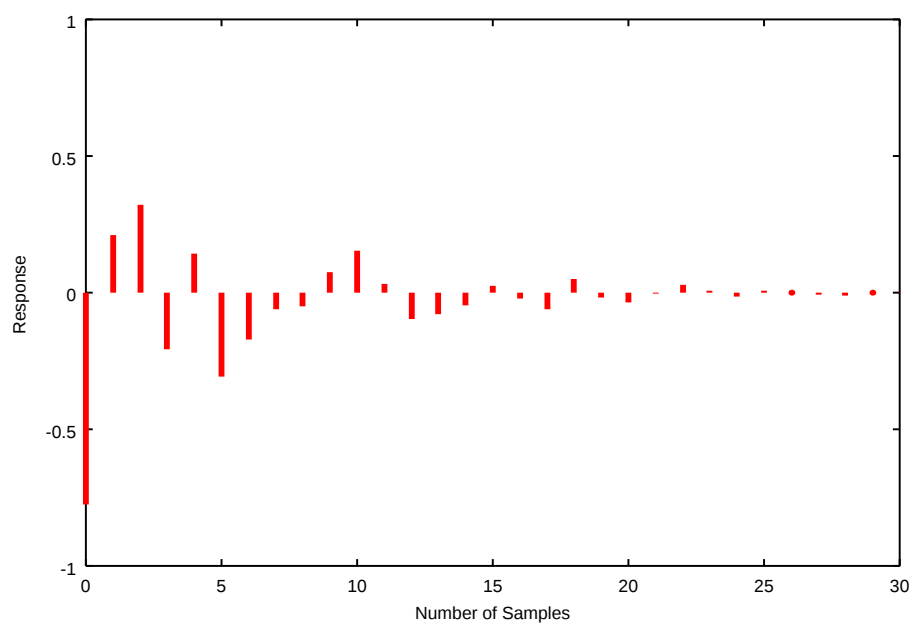


図 5.5 左側のインパルス応答

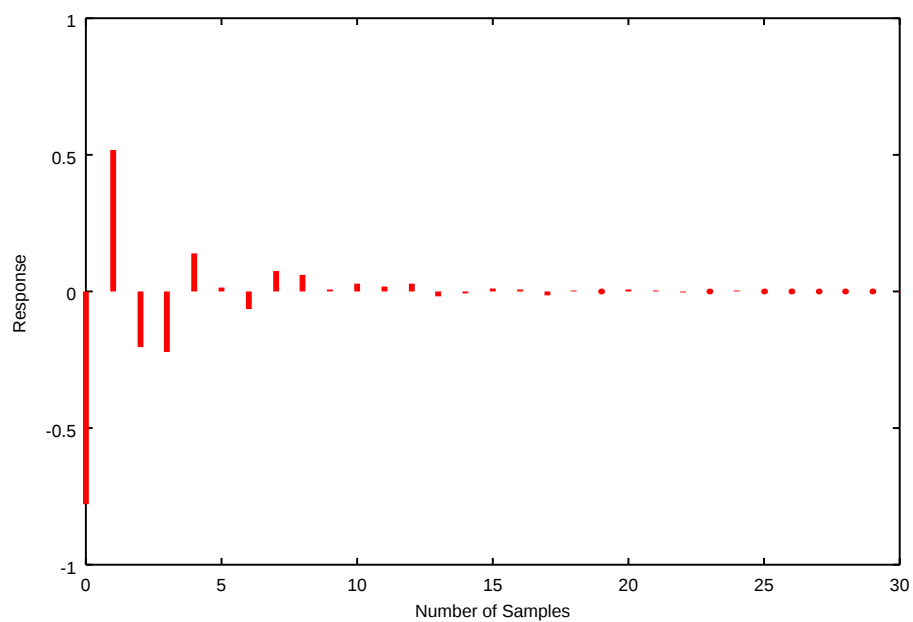


図 5.6 右側のインパルス応答

5.4 まとめ

クロストーク成分が大きく違う環境

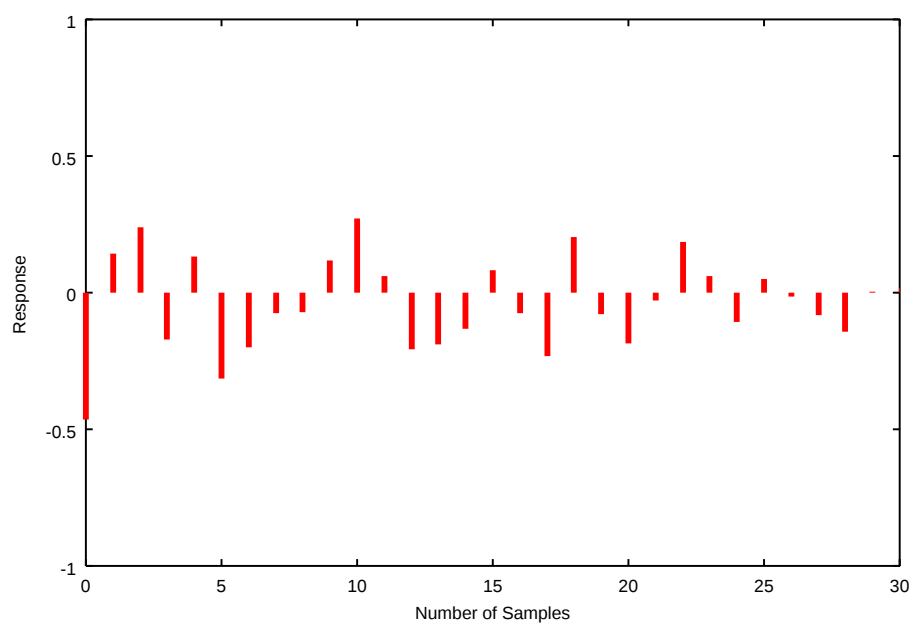


図 5.7 左側のインパルス応答

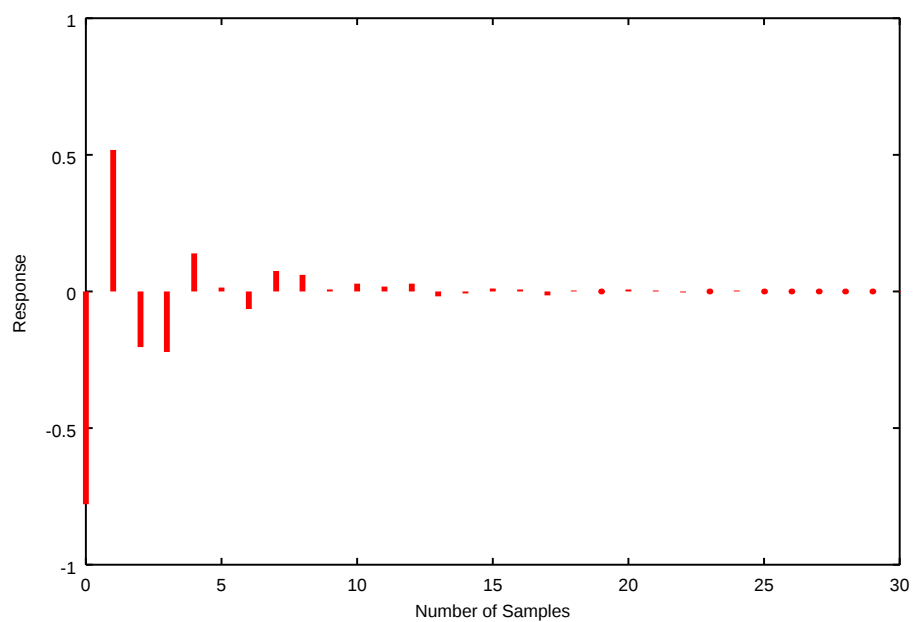


図 5.8 右側のインパルス応答

5.4 まとめ

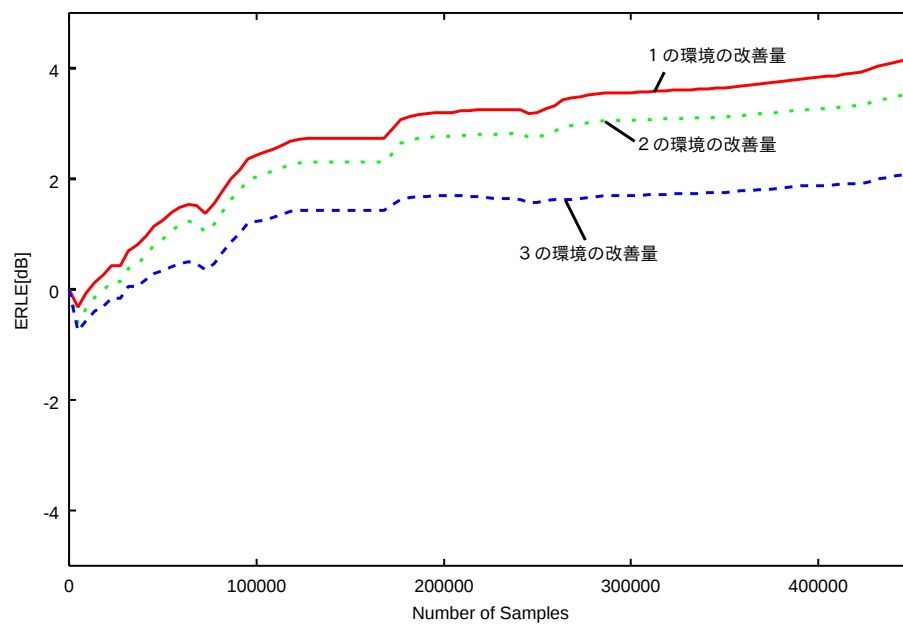


図 5.9 それぞれの環境での改善量の収束特性

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ていたが，これは受聴者が動く等によって伝達特性が変化しないと仮定した場合でのみ有効性が保証されている．そこで，受聴者が移動したことなどによって伝達特性が変化した環境を想定した室内インパルス応答を用いて多入力信号補正システムの性能評価を行った．

計算機シミュレーションにより，左右のクロストーク成分の違いが大きくなるほど改善量も減少していく傾向が確認された．これによって，左右のクロストーク成分に相関がない場合システムとして期待される信号の補正効果が減少することが検証された．

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クロストーク成分の違いによる改善量の傾向を示した．しかし，今回シミュレーションに用いた伝達関数は乱数によって変化をつけたものであるため，実際の環境において伝達関数に違いが出ると改善量にどの程度の影響を与えるか定量的に評価する必要がある．また，伝達関数の変化による改善量の低下が起こった場合の対策案が必要になってくる．

謝辞

本研究において、夜遅くまで御指導いただいた福本 昌弘教授に深く感謝致します。研究以外でもこれから社会に出る上で大変多くのためになることを学ばさせていただきました。また、本研究を審議してくださった吉田 真一講師，ならびに情報システム工学科教員の皆様に深く感謝致します。遅ればせながら提出した梗概に対しいただいた吉田講師の御助言はためになり、やる気を促進されました。発表の資料の添削を何度も行っていたいただいた福富 英次助手にも感謝致します。食べ物を御馳走になったり、また、輪講を始め様々な知識を学ぶ機会をくださった佐伯 幸郎助手にも御礼申し上げます。お忙しい中、小さなことでも快く丁寧に質問に答えてくださった金井 宏一郎氏にも感謝致します。ポン菓子をいただき、研究に臨む上での活力をいただいた徳久 翔太氏にも御礼申し上げます。

同研究室で共に研究生活を乗り越えた上岡 良太氏にも感謝致します。同期の中でどんなことも積極的にこなしていき頼りになるリーダーシップを発揮し大変見本になりました。社会人になり車を持つようになれば「その道」で会うかもしれません。そのときはどうぞよろしくお願いします。同じく同研究室同期の森田 徹氏にも感謝致します。苦難をも笑いに変え、前へと進む大切さを学び、そして気力を湧かせるに至りました。その接しやすい人間性は是非とも見習っていきたいと思います。また、同研究室同期の船橋 尚樹氏にも感謝致します。特定の分野にかける情熱にはいろいろと学ばさせていただくことができました。同研究室同期、島田 卓哉氏にも感謝致します。高校からのつき合いで大学生活で共にする時間はもしかしたら一番多かったかもしれません。大学生活もですが、高校から7年に至る人生の中で、苦楽を共にしてこれたことにこの場をお借りして深く御礼申し上げます。

また、吉田研究室 鈴江 直人氏、高田研究室 加藤 孝氏には、共に研究生活を励ましあいながら今日までこられたことに深く感謝致します。

最後に、公私ともに大学生活において支えてくださった全ての皆様に御礼申し上げます。

参考文献

- [1] 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”, 信学技法 SIP2004-115, vol.104, no.559, pp.31-36, 2005.
- [2] 豊永雅行, 福本昌弘, “多入力信号補正による収束特性の評価法”, 平成 16 年度 学士学位論文, 2005.
- [3] 三上直樹, “C 言語によるデジタル信号処理入門”, CQ 出版社, 2002.
- [4] S. ヘイキン (武部 幹 訳), 適応フィルタ入門, 現代工学社, 1994.
- [5] 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.
- [6] 北脇信彦, デジタル音声・オーディオ技術, 電気通信協会, 1999.
- [7] M.Miyoshi and Y.Kaneda, “Inverse filtering of room acoustics”, IEEE Transactions on Acoustics, Speech, and Signal Processing, 1988.
- [8] 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.