

平成 19 年度
学士学位論文

ネットワーク上のリソースを活用した実時間音場再生システムの構築

Construction of a real time sound field reproduction
system which utilized the resource on a network

1080423 山田 良

指導教員 福本 昌弘

2008 年 3 月 7 日

高知工科大学 情報システム工学科

要 旨

ネットワーク上のリソースを活用した実時間音場再生システム の構築

山田 良

近年，音響の分野では音に臨場感を与える手法として原音場の再現による音場再生が注目されている．この音場再生は実時間で行うことが望まれており，デジタル信号処理専用プロセッサを用いることで実現可能となっている．一方，安価に高性能なマイクロプロセッサ導入できることから汎用コンピュータは高性能な処理装置を備えている．

本研究ではネットワーク上に存在する複数の信号処理資源を活用することで，分散処理可能な実時間音場再生システムの構築を行っている．従来手法の音の入出力と汎用コンピュータとの違いを検証し，適応信号処理を音の入出力に対応させている．有効性の検証は実験によって行い，提案手法が２台の高性能な機器による分散処理を行ったとき最も処理性能が向上することが明らかになった．しかし，実験で音場再生の安定は難しく，その原因として遅延量を確認を行ったところ処理負荷が大きいとき遅延量の変化が起きやすくなり同期を取ることが非常に困難となることが明らかになった．また実験環境以外で音場再生を実行したとき考えられる問題点を考察して処理に必要な時間を変化させることによる解決方法を提案している．

キーワード 音場再生 実時間処理 分散処理

Abstract

Construction of a real time sound field reproduction system which utilized the resource on a network

Ryo Yamada

In recent years, the sound field reproduction by reappearance of an original sound field attracts attention as the technique of giving presence to sound in acoustical society. The sound field reproduction that is hoped to perform actual time. Sound field development of sound field reproduction system using Digital Signal Processor. On the other hand, general-purpose computer has high-performance processing unit from the ability to introduce a cheaply highly efficient microprocessor.

In this research, the realtime sound field reproduction system in which distributed processing is possible is built by utilizing two or more signal-processing resources which exist on a network. The difference from input and output and the general purpose computer of the sound of the conventional technique is verified. Adaptive signal processing is made to correspond to input and output of sound. The experiment is performing the check of validity. When the proposal technique performed the distributed processing by two highly efficient apparatus, the processing performance improved most. However, it was difficult to stabilize sound field reproduction in an experiment. When the amount of delay was checked as a cause, and processing load was large, it became clear that it delay to take place easily, and to take a synchronization. Moreover, the solution by changing time required for processing in consideration of the problem considered when sound field reproduction is performed except experiment environment is proposed.

key words sound field reproduction,real-time proceccing,decentralized processing

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	3
1.3	本論文の構成	4
第 2 章	音場再生システムと逆特性の近似	5
2.1	はじめに	5
2.2	音響システムと線形モデル	5
2.3	適応フィルタを用いた逆特性の近似	7
2.4	多チャンネル音場再生システムにおける逆特性近似	9
2.5	まとめ	10
第 3 章	多入力多出力システム	11
3.1	はじめに	11
3.2	2 チャンネル-2 点制御	11
3.3	多入力信号補正法	12
3.3.1	逆特性の近似的推定法	12
3.4	適応フィルタの更新	15
3.5	まとめ	15
第 4 章	実時間音場再現システムの構築と評価	17
4.1	はじめに	17
4.2	提案システムの構成	17
4.2.1	サウンドデバイスの制御	18
4.2.2	音場再生システムの実装	19

目次

4.2.3	負荷の分散方法	20
4.3	実験による評価	20
4.3.1	実験条件	21
4.3.2	実験結果と考察	22
4.3.3	提案システムの問題点	23
4.4	WAN 上での問題点と解決案	24
4.4.1	データの欠落による影響	24
4.4.2	ネットワーク遅延の解決案	26
4.5	まとめ	27
第 5 章	結論	28
5.1	本研究のまとめ	28
5.2	今後の課題	28
謝辞		29
参考文献		30
付録 A	広い領域での音場再生法	31

図目次

1.1	バイノーラルシステムの例	2
1.2	トランスオーラルシステムの例	2
1.3	MINT 理論によるトランスオーラルシステム	3
2.1	閉空間における信号伝搬の様子	6
2.2	図 2.1 の等価回路	7
2.3	室内伝達特性の補正	7
2.4	MINT 理論に基づく音場制御系	8
2.5	MINT 理論に基づく音場制御系	9
3.1	2 チャネル-2 点制御系モデル	12
3.2	2 チャネル-2 点制御系モデル (クロストーク無し)	13
3.3	逆近似フィルタと構成図	14
4.1	システムの全体構成	18
4.2	内部バッファ	18
4.3	実装した音場再生システム	19
4.4	負荷分散手法	21
4.5	遅延の確認	24
4.6	補正サーバとクライアントの WAN 接続	25
4.7	補正サーバと補正サーバの WAN 接続	26
A.1	境界音制御の概念図	31

表目次

4.1	使用した機器	22
4.2	実時間処理可能なタップ数	23

第 1 章

序論

1.1 はじめに

近年，ネットワークインフラ環境の整備に伴い，様々なコンテンツがブロードバンドで伝送することが可能になった．IP 電話やネットワークミーティングなど，今まで別の伝送路を用いて伝送されていたものもネットワーク上で実現されている．既存の技術がネットワーク上に置き換わり，活用範囲が広がりつつある．そういった，ネットワークの利用者にとって，超高解像度の HDV 映像やマルチチャネルによる音響再生等といった，マルチメディアに関するコンテンツは関心の高い分野の一つである．現在では，光ファイバの利用者が増加し高解像度の映像や 5.1ch のサラウンド再生方式といったマルチチャネルのデータを伝送することも可能である．しかし，高音質化だけでは録音した空間と同じ臨場感を得ることが出来ないことが知られている．

原音場と同じ臨場感を得る手法は大きく 2 つに分類される．1 つは映像に対して音像移動を同期させることで実現するもので，もう 1 つは録音した空間と同じ音場を再現させる手法である．前者は，原音場の特性を物理的に再現し，受聴者に臨場感が伴うように再生音場の特徴を合成させるものである．これは主に映像に合わせて音が到達する合成加工を目的とした技術である．映像に対する音像定位を実現するには，音圧レベルや到達時間の差また位相の差をコントロールする必要がある．ホームシアター等の 5.1ch 再生方式はこの手法を取っている．しかし，この手法では再生環境については考慮されてないため，無響室等の再生環境による影響が少ない部屋でのみ実現可能となっている．また，再現範囲が広い場合再生するスピーカ数が増加するが，制御の複雑になることから，正確に再現することは困難であり

1.1 はじめに

実用化はされていない。

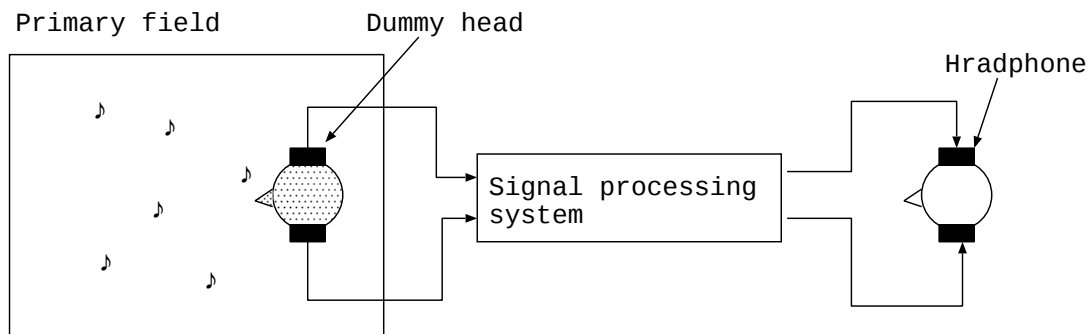


図 1.1 バイノーラルシステムの例

後者は、原音場の特性を受聴点において忠実に再現する手法である。原音場を忠実に再現するためには、再生音場での反射や残響と言った影響を考慮して再現する必要がある。それらの影響を最小限にしたシステムにバイノーラルシステム 1.1 がある。ダミーヘッドと呼ばれる人の頭部を模したマイクで收音することで、頭部や耳の形状による影響を受けた收音データを得ることができる。さらに、再生する空間を受聴者の耳元に限定することにより、簡単な構成で録音環境を再現することが可能である。しかし、ダミーヘッドと受聴者の形状が異なる場合正確に音場を再現することができない。また、ヘッドフォン再生することでヘッドファンの周波数特性の影響を受ける。さらに、ダミーヘッドと受聴者の耳の形状による影響を 2 度受けるため音像定位の劣化がある。バイノーラル收音された原信号を、ラウ

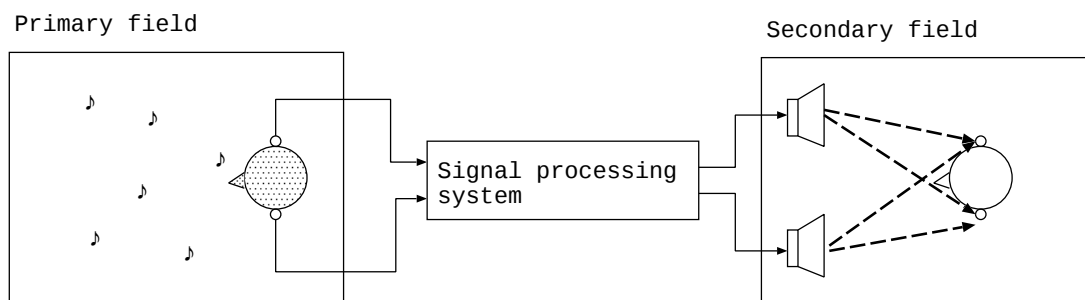


図 1.2 トランスオーラルシステムの例

ドスピーカによる再現手法に関しても研究が進められている。スピーカから再生され空間という伝達経路の影響を受けた信号を受聴点において再現するために逆フィルタを用いたシ

1.2 従来手法の問題点

ステム図 1.2 をトランスオーラルシステムという．このシステムは，MINT 理論に基づき受音点以上のスピーカを用いることでトランスオーラルシステムが実現可能である．しかしながら，MINT 理論によって構成されたシステムは受音点の数が増加すると，ラウドスピーカ数が増加し制御の複雑化やコストの増加，逆フィルタの近似に必要な演算コストとメモリ量の増加などがある．もし，残響時間の長い環境でトランスオーラルシステムを構築した場合，逆フィルタのタップ数が非常に長く実時間で動作できないシステムになるという問題もある．

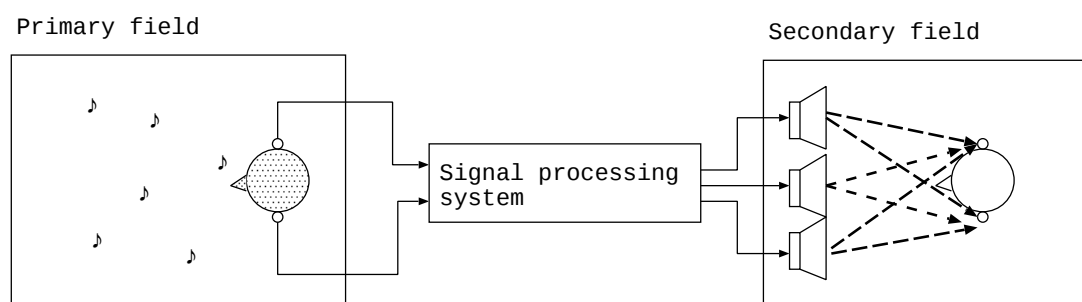


図 1.3 MINT 理論によるトランスオーラルシステム

1.2 従来手法の問題点

以上述べたように，トランスオーラルシステムでは再生環境の変化に対する追従性能と残響や反射の影響を補正する演算に関したコスト，構成システム規模の増大が挙げられた．実時間システムでは演算コストの増加は再生音の連続性に影響を及ぼしてしまう．音が連続的に再生されない場合，音は途切れ途切れに再生される．この時，逆システムの推定において入力に対する結果が全く異なるしまうため正しく環境による影響を取り除くことができない，

使用者の環境として大規模なハードウェア構成と要するシステムは現実的ではない．そのため，利用環境を選ばないシステムを構成する必要がある．これらの問題に対して，ハードウェア構成を見直した，2 個のラウドスピーカを用いて 2 点の制御点に対して補正を行う信号補正システムが考案されている，これは，スピーカから制御点へ交差して到達するクロス

1.3 本論文の構成

トーク成分を再生環境の影響に加えて逆フィルタを設計することでハードウェアの簡素化が可能になった．逆フィルタの算出には多くの演算時間が必要なことが知られている，また，環境の変化に追従するには収束特性の優れたアルゴリズムを使用する必要がある．より実用的なシステムにするには優れたアルゴリズムの演算時間を実時間で可能にする必要がある．本研究では多くの処理時間の必要な逆システム推定および補正を複数サーバによる分散処理で実時間可能なシステムの構築を目的とする．

1.3 本論文の構成

本論文の構成について述べる．

第 2 章では 基本的な技術である適応アルゴリズムについて述べる．また，MINT 理論による多チャネル音場再生技術における従来の逆フィルタ推定法を紹介する．

第 3 章では 本研究で用いられている 2 チャネル-2 点制御による音場再生システムについて述べ，実時間処理可能な分散システムの構築手法を提案する．

第 4 章では，実験によって得た結果から，従来手法と提案手法について考察する．

最後に，以上で得られた結果を元第 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は，システムを構成するハードウェアや再生手法の違いから様々なシステムが考案されてきた．本章では，音場などの環境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．次にラウドスピーカを用いた音場再生手法の考え方と，実現方法として逆フィルタを用いる手法について説明を行い，多入力多出力の入出力系における逆フィルタの設計について説明する．

2.2 音響システムと線形モデル

空間中の 2 点（音源と受音点など）の音の伝わり方を示す関数を空間伝達関数 $G(z)$ という．伝達関数 $G(z)$ は空間が屋外などの自由空間の場合，2 点間の距離のみに依存する．これは音圧が音源までの距離に反比例して減衰する特性と，その距離を音が伝搬するのに要する伝搬遅延時間によるものである．一方，音場が図 2.1 に表すような反射性を持つ閉空間において，音源であるラウドスピーカと受音点であるマイクロフォンを，2 点間での音の伝搬について考える．ここでラウドスピーカに入力した音を $x(t)$ ，マイクロフォンで受音した信号を $y(t)$ とする．このときマイクロフォンで観測される信号 $y(t)$ には，音源から受音点に直接到達する直接音と，壁面で反射された多数の反射音が含まれる．反射音はマイクロフォンで受音されるまでに壁面で反射された回数によって 1 次反射，2 次反射などに分類される．また，音の反射は空間による減衰と壁面による吸音などによって音のエネルギーが消滅して

2.2 音響システムと線形モデル

しまうまで繰り返される．音源からパルス波が出力されたとき，マイクロフォンで観測された信号はインパルス応答 と呼ばれる．この観測信号は時間の経過と共に直接音が観測される

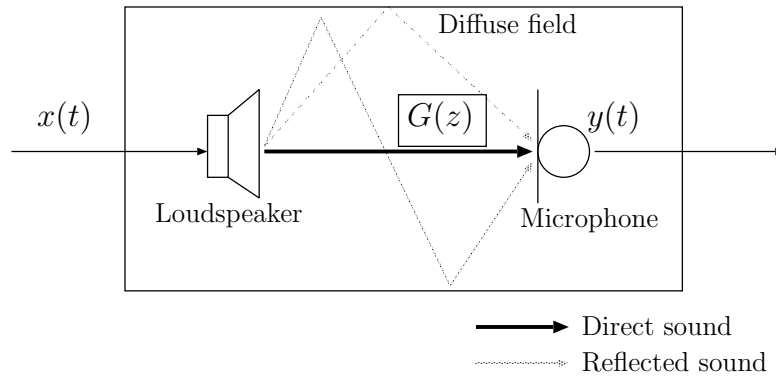


図 2.1 閉空間における信号伝搬の様子

次に一次反射音，二次反射音などの低次反射音が散発的に観測される．その後，続けて高次反射音が観測される．ここで，低次の反射音は単に初期反射音と呼ばれ，高次の反射音はインパルス応答長が時間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本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，マイクロフォン，空間音響伝達系による影響が全て線形であると仮定すると．入力信号 $x(t)$ と出力信号 $y(t)$ の関係はインパルス応答 $g(k)$ と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 $y(n)$ が求められることを意味している．両辺を z 変換することで次式を得ることができる．

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

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 の関係から，図 2.1 伝達系は図 2.2 のような等価回路として表すことができる．このように室内音場におけ

2.3 適応フィルタを用いた逆特性の近似

る音響システムは線形モデル化することができる．

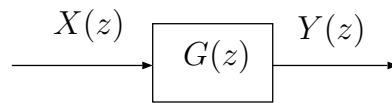


図 2.2 図 2.1 の等価回路

2.3 適応フィルタを用いた逆特性の近似

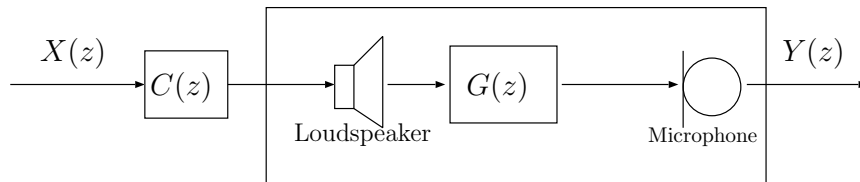


図 2.3 室内伝達特性の補正

トランスオーラルシステムを用いた音場再生では，ラウドスピーカに入力される所望信号が受聴点での観測信号と等しくなるように制御することを目的としている．つまり図 2.2 に示す線形モデルにおいて， $X(z) = Y(z)$ となるように制御する必要がある．しかし実空間上の音場では 2.3 で示したとおり，受聴点における観測信号は伝達特性 $G(z)$ の影響を受けるため，この影響を補正する必要がある．そこで，2.3 に示すように所望信号に対して補正フィルタ $C(z)$ を用いて制御を行う．この補正フィルタは空間伝達特性の逆特性を近似したフィルタである．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と仮定すると，

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

という関係が成り立つ．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で

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

2.3 適応フィルタを用いた逆特性の近似

となる．したがって伝達特性が 1 を示すとき，受聴点では所望信号 $x(t)$ がそのまま観測されることになる．ただしこのような関係が成り立つためには伝達関数 $G(s)$ の零点が単位円上に存在する最小位相系でなければならない．しかし一般的に室内伝達特性は非最小位相系であり，正確な逆特性を推定する事は困難であることから，式 2.5 のような関係は実現できない．

非最小位相系で逆近似を実現するための方法として，逆特性の誤差エネルギーを最小にすることで伝達特性の近似的に求める手法が提案されている．代表的な手法として FIR フィルタを用いて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) を最小にする $C(z)$ の探索問題に帰着する．ここで $G(z)$ が最小位相系となる場合，逆特性 $1/G(z)$ の近似は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ことから，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の除去や，遅延や雑音を付加による方法が提案されている．

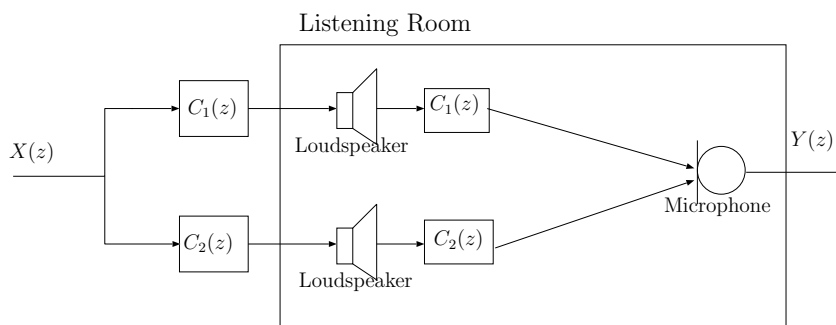


図 2.4 MINT 理論に基づく音場制御系

一方，MINT の理論によると，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について示す．受音点において観測信号 $Y(z)$ は

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

2.4 多チャンネル音場再生システムにおける逆特性近似

と表される．この系において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 $C_1(z)$ ， $C_2(z)$ を求める必要がある．

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

この解の存在条件として， $G_1(z)$ ， $G_2(z)$ が有限であり，かつ共通の零点を持たないという条件がある．この条件を満たすとき，式 2.8 を満たす $C_1(z)$ ， $C_2(z)$ が存在することを数学的に証明されている．

2.4 多チャンネル音場再生システムにおける逆特性近似

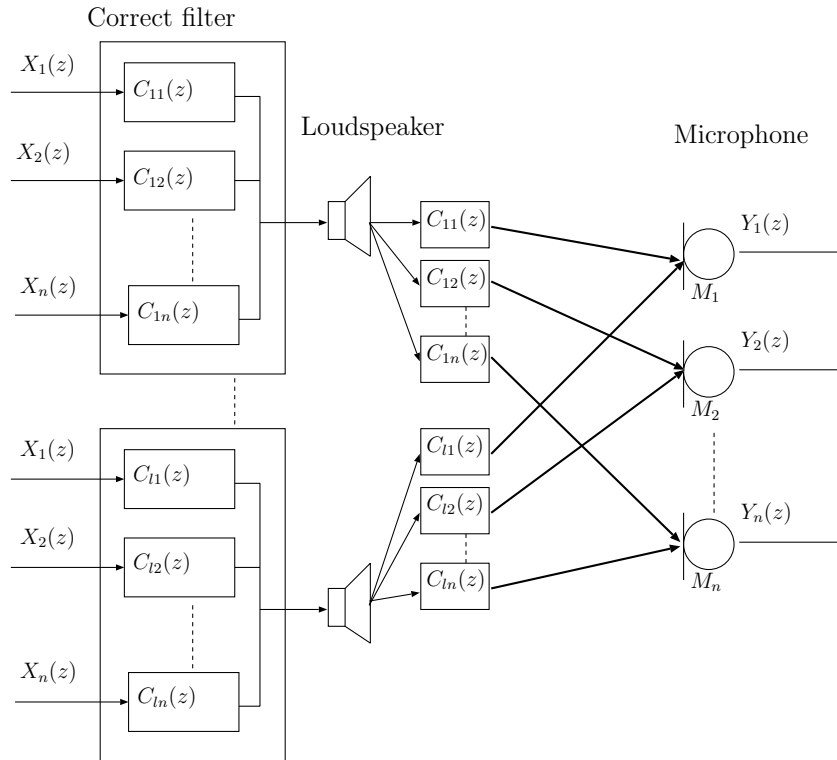


図 2.5 MINT 理論に基づく音場制御系

図 2.5 に音減数 l ，制御点数 n 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の構成を示す．室内伝達関数のインパルス応答を $G_{jk}(z)$ ，補正フィルタを $C_{jk}(z)$ ，所望信号を $X_l(j)$ ，観測信号を $Y_l(z)$ とする．また $j(= 1, 2, \dots, l)$ および $k(= 1, 2, \dots, n)$ はそれぞれ音源及び制御点の番号である．多チャンネル音場再生システムで

2.5 まとめ

は，各制御点に対する制御を他の点とは独立に行うことが望ましい．そこで補正フィルタ群においては入力信号 $X_j(z)$ をパルスとしたときに，制御点 M_j に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

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

を満たす $C(z)$ を求めればよい．ただし

$$G(z) = \begin{bmatrix} G_{11}(z) & G_{12}(z) & \cdots & G_{1l}(z) \\ G_{21}(z) & G_{22}(z) & \cdots & G_{2l}(z) \\ \vdots & \vdots & \ddots & \vdots \\ G_{n1}(z) & G_{n2}(z) & \cdots & G_{nl}(z) \end{bmatrix} \quad (2.10)$$

$$C(z) = \begin{bmatrix} C_{11}(z) & C_{12}(z) & \cdots & C_{1l}(z) \\ C_{21}(z) & C_{22}(z) & \cdots & C_{2l}(z) \\ \vdots & \vdots & \ddots & \vdots \\ C_{n1}(z) & C_{n2}(z) & \cdots & C_{nl}(z) \end{bmatrix} \quad (2.11)$$

とし， I_l は $l \times l$ 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，それらの 2 次音源毎に畳み込むことで逆特性を近似する制御系が構築される．

2.5 まとめ

本章では音場再生システムが線形モデルで表されることについて説明し，伝達特性の影響に関して補正を行う近似逆フィルタが設計できることを示した．また，MINT 理論を用いることで多チャネル-多点制御系での音場再生が可能であることを述べた．しかし一般的な室内環境を再現空間としたとき，残響時間が長くなる場合がほとんどである．つまり，空間伝達特性のインパルス応答長が長くなることから演算量の増加という問題が起こるため，MINT 理論を用いた音場再生システムは実用的ではない．第三章ではこの問題を解決するための考案された手法について述べる．

第 3 章

多入力多出力システム

3.1 はじめに

ラウドスピーカによる多チャンネル音場再生システムを用いて室内伝達特性の影響を除去するには近似逆フィルタをが不可欠である．そのための方法として，MINT 理論に基づいた近似逆フィルタの設計手法が提案されている．しかし前章において空間伝達関数の残響時間が長くなる環境では，計算量と要求されるメモリ量が膨大となるため実用的なシステムには向いていないことを説明した．これに対して，2 つ制御点と 2 つのスピーカで音場再生を行う多入力信号補正システムが考案された．本章ではこの多入力信号補正システムについて解説する．

3.2 2 チャンネル-2 点制御

ここでは 2 つの音源を用いて 2 つの制御点に対して制御を行う音響系について線形モデル化を行う．このモデル化の結果から 2 つのスピーカから出力される原音と制御点での観測信号が一様に等価できないことを示す．図 3.1 に 2 チャンネル-2 点制御系モデルを示す．観測点における観測信号 y_i は入力信号 x_i および室内伝達特性 $g_{ij}(k)$ を用いて以下の式で与えられる．

$$y_1(t) = x_1(t)g_{11} + x_2(t)g_{21} \quad (3.1)$$

$$y_2(t) = x_2(t)g_{22} + x_1(t)g_{12} \quad (3.2)$$

3.3 多入力信号補正法

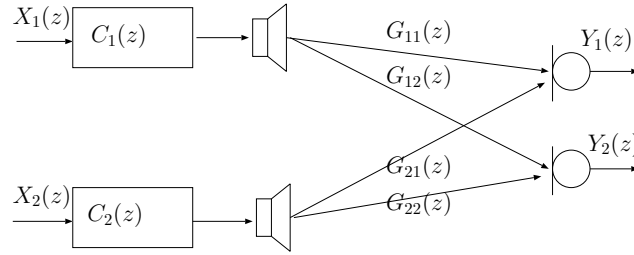


図 3.1 2 チャネル-2 点制御系モデル

ここで式 3.2 における観測信号は，入力信号 x_1 が伝達特性 g_{11} の影響を受けて届く直接信号と，入力信号 x_2 が伝達特性 g_{21} の影響を受けて届くクロストーク信号の総和となる．また，式 3.2 においても同様に直接信号にクロストーク信号の総和となる．クロストーク信号とは，ある受聴点に伝わるはずの信号が別の受聴点に直接伝わってしまう信号のことである．以上のことから補正フィルタ $C_i(z)$ は直接信号の伝達特性の影響とクロストークの影響を除去あるいは軽減できることが望まれる．

3.3 多入力信号補正法

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手法について説明をする．

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手法について述べる．説明簡略化のため，ここでは図 3.2 クロストークの影響がない線形モデルについて考える．原音である所望信号 $x_i(t)$ が補正フィルタ $C_i(t)$ 通過後の信号 $d_i(t)$ は次のように表される．

$$d_i(t) = c_{i,L} x_{i,L}^T(t) \quad (3.3)$$

ここで

$$c_{i,L} = [c_{i,0}, c_{i,2}, \dots, c_{i,L}] \quad (3.4)$$

3.3 多入力信号補正法

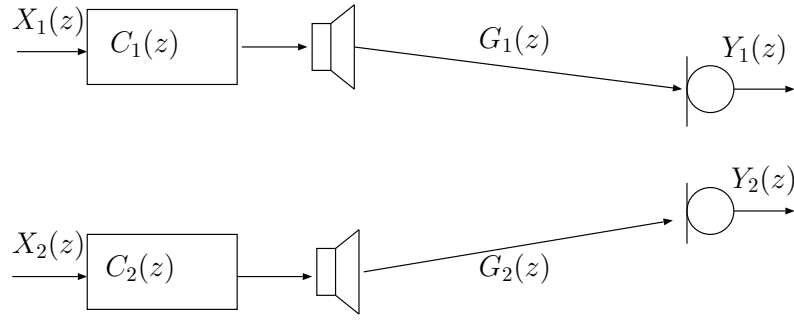


図 3.2 2 チャンネル-2 点制御系モデル (クロストーク無し)

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

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

$$y_j(t) = d_{i,L_g}(t) g_{ij,L_g}^T \quad (3.6)$$

となる．ただし

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

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

である．ここで補正フィルタ $C(z)$ が伝達関数 $G(z)$ に対する近似逆フィルタと同様の性質だと仮定すると， z 変換による表現により以下に示される近似式が成り立つ．

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

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が等しくな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 $h_{ij,L}(t)$ ，適応フィルタの入力状態ベクトルを $y_{ij,L}(t)$ とすると，入出力関係は

$$s_{ij,L}(t) = h_{ij,L}^T(t) y_{ij,L}(t) \quad (3.10)$$

3.3 多入力信号補正法

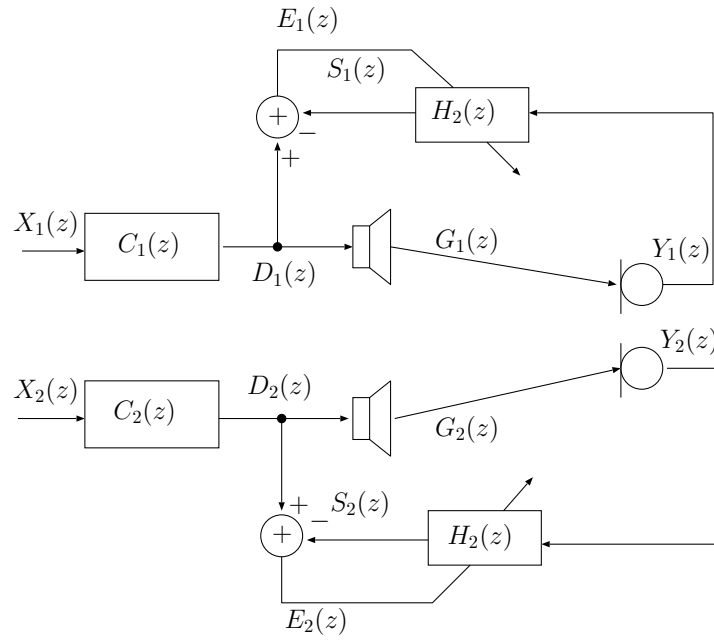


図 3.3 逆近似フィルタと構成図

で与えられる．ただし

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

$$y_{j,L}(t) = [y_j(t), y_{t-1}(t-1), \dots, y_j(t-L)] \quad (3.12)$$

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 $e(t) = 0$ の場合を考えると

$$y_{j,L}(t) h_{ij,L}^T(t) = x_{i,L}(t) c_{j,L}^T(t) \quad (3.14)$$

となる．式 3.14 において適応フィルタ係数 $h_{ij,L}$ を補正フィルタ係数 $c_{ij,L}$ として与えると，

$$\begin{aligned} y_{j,L}(t) h_{ij,L}^T(t) &= x_{j,L}(t) h_{ij,L}^T(t) \\ y_{j,L}(t) &= x_{i,L}(t) \end{aligned} \quad (3.15)$$

となる．これより，入力誤差 $e(t) = 0$ でなるならば，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 $e_{ij}(t)$ が最小になるように適応フィルタ

3.4 適応フィルタの更新

の係数を更新していく必要がある．そこで係数更新には

$$h_{ij,L}(t+1) = h_{ij,L}(t) + \alpha \frac{y_{j,L}(t)}{\|y_{j,L}(t)\|^2} e_{ij}(k) \quad (3.16)$$

で示される学習同定法を用いる．ここで、 $y_{j,L}(t)$ は入力状態ベクトル、 $h_{ij,L}(t)$ 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、フィルタ係数の修正量を制御するパラメータである．

3.4 適応フィルタの更新

補正フィルタの係数を求めるために、学習同定法に基づいて適応フィルタを更新する．次式を入力信号の 1 サンプルデータ毎に繰り返し処理を行うことにより、右側音原が伝達特性 G_{11} 影響を受け通過する信号が所望信号である $x_1(t)$ に近似していくと考えられる．もう一方の入力信号に関しても同様のことが言える．まず、2 つの近似逆フィルタ $H_{ij}(z)$ に対する出力誤差 $e_{ij}(t)$ をそれぞれ以下のように計算する．

$$e_{ij}(t) = d_i(t) - s_{ij}(t) \quad (3.17)$$

次に各逆フィルタ $H_{ij}(z)$ の係数を学習同定法を用いて以下のように算出する．

$$h_{ij,L}(t+1) = h_{ij,L}(t) + \alpha \frac{y_{j,L}(t)}{\|y_{j,L}(t)\|^2} e_{ij}(k) \quad (3.18)$$

ここで、 α はステップゲインである．左右の所望信号に対する補正フィルタ係数 c_i をそれぞれ以下のように更新する．

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

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

3.5 まとめ

本章では、2 チャネル-2 点制御による原音場再生システムを紹介した．このシステムはクロストーク成分の相互相関に着目し、必要とするハードウェア構成および計算量を抑えた音

3.5 まとめ

場再現手法である．また学習同定法に基づいた補正フィルタ係数の算出法および係数の更新手順についてまとめた．

第 4 章

実時間音場再現システムの構築と評価

4.1 はじめに

本研究では再現空間における機器の処理負荷軽減を目的としており，利用者が必要とする再生機器の処理能力に依存しない実時間音場再生システムを構築を目指している．本章では，始めに提案したシステムの全体構成を示す．次に実時間で音場再生を行う場合に利用されている DSP(Digital Signal Processor) と本システムで用いているサウンドデバイスとの音の入出力の違いを説明する．そして，音の入出力を考慮した音場再生システムの実装手法と分散処理の手順を述べる．最後に本手法によって構築したシステムの有効性を実験を行うことにより確認する．また，実験環境以外で本システムを使用した場合について考察を行う．

4.2 提案システムの構成

全体構成を図 4.1 に示す．利用者はネットワークを経由し，録音空間で録音・補正を行っている補正サーバにアクセスすることでこのシステムを使用することができる．このシステムにおいてクライアントマシンはサーバから送信される再生音を再現空間でスピーカから再生し，また同時に受聴点に設置したマイクで録音をする機能が必要となっている．そして再生した原信号と録音した観測音を補正サーバに転送を行う．一方，この 2 つの信号を受け取った補正サーバはさらに推定サーバへ転送する．これにより補正サーバは推定サーバともに負

4.2 提案システムの構成

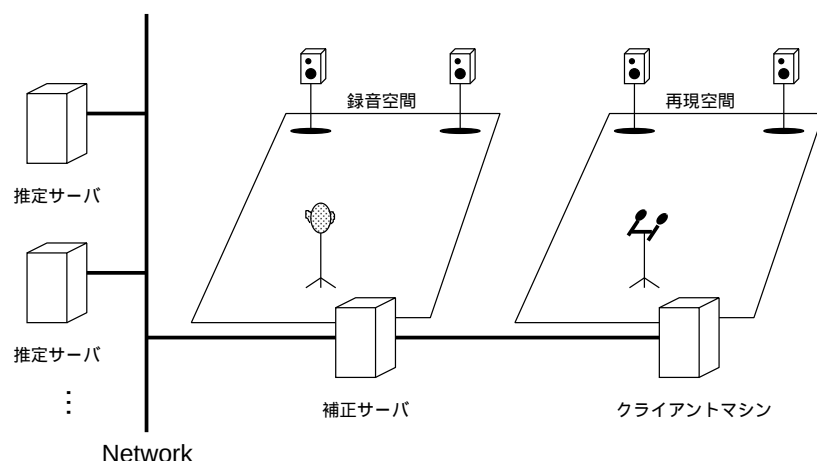


図 4.1 システムの全体構成

荷を分散させて動作することでより高性能な音場再生システムを構築することができる。

4.2.1 サウンドデバイスの制御

実時間で音場再生を行う手法として DSP によって信号補正システムを構築する研究がある [5]。この研究では録音および再生信号の入出力は割り込みとして処理されており、標準化周波数に基づいた一定間隔内に補正処理と推定処理が完了している場合実時間で音場再生が可能になっている。しかし、DSP は一般に普及していないことから本研究の目的としている音場再生システムには適していない。よって本研究ではコンピュータに音の入出力機能を拡張するサウンドカードを用いることにする。また、このデバイスを OSS(Open Sound System) によって制御する。



図 4.2 内部バッファ

サウンドカードは図 4.2 に示されるような内部バッファを持ち、fragment と呼ばれるブロック単位により分割管理されている。そのことから、OSS では以下に示す条件により信号の入出力を行う必要がある。

4.2 提案システムの構成

- 録音・再生は内部バッファが fragment 分の信号を持っているとき可能である．
- バッファが完全に使用されているときデータの書き込みがブロックされる．
- 録音されたデータで満たされた fragment がないと読み込みがブロックされる．

よって，本システムでは従来手法と異なり fragsize のデータが録音・再生される一定間隔で処理を行う必要がある．この一定間隔中に伝達特性の補正，逆特性の推定，補正フィルタの更新処理を行う．

4.2.2 音場再生システムの実装

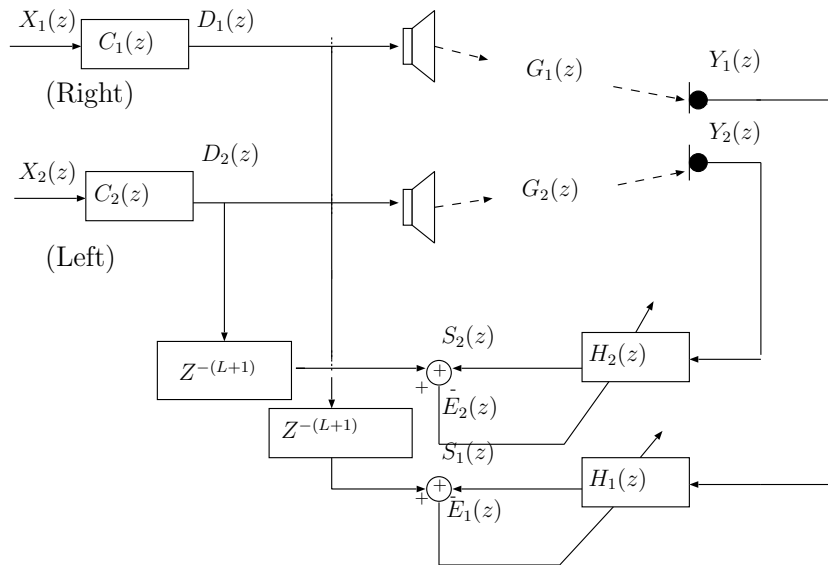


図 4.3 実装した音場再生システム

提案システムにおいて実装している適応信号処理システムのブロック図を 4.3 に示す．1 チャンネル-1 点制御系をステレオで動作させたもので，問題を簡単にするためにクロストークの影響を考慮していない．ここでは，単位遅延器 Z^{-1} を縦続接続することで再生環境のインパルス応答長 $L + 1$ の遅延を与えている．適応信号処理は因果律により未来の信号を推定できない．つまり，スピーカから再生した瞬間の再生信号に比べ録音した観測信号はスピーカからマイク間の空間を伝わるまでに伝達特性のインパルス応答長 $L + 1$ の時間を必要とし

4.3 実験による評価

ており，その遅延量 $L + 1$ だけ未来の信号となる．このことから再生信号に $L + 1$ の遅延を与えることで因果律を満たすことができる．

また，信号補正手法は 3.3 で述べた逆特性の推定手法により fragment サイズの信号を含んでいるブロック単位で行っている．このブロック単位で本システムを動作させているため，再生信号はブロック毎に同じフィルタ係数で補正処理を行っている．そして，観測信号も再生環境の伝達特性の影響を受けたブロックデータとなる．この再生信号と観測信号から次の逆特性の推定を行うにはブロック当たり 1 回のフィルタ係数の更新が可能である．

4.2.3 負荷の分散方法

本システムでは以下の手順によって補正サーバが推定サーバに必要としているパラメータを与えることで，ネットワーク上の資源を活用した負荷分散を行っている．

1. クライアントから再生音と観測音を受け取る．
2. 受け取ったデータを推定サーバへ転送する．
3. 推定サーバでは受け取ったパラメータから逆特性の計算を行う．
4. フィルタ係数更新タイミングで補正サーバにフィルタ係数を適応する．

さらに複数の機器を用いての負荷分散手法は図 4.4 に示すように，更新手順 (2) のときに N 台の推定サーバに再生・観測信号をパラメータとして転送し，(4) で受け取るフィルタ係数を補正サーバで分割して処理することで可能にしている．

4.3 実験による評価

本研究により構築したシステムの有効性を検証するために，実際に録音空間と再現空間となる高知工科大学 A502, A503 に各処理装置とスピーカ，そしてマイクロフォンを設置して実験を行う．

4.3 実験による評価

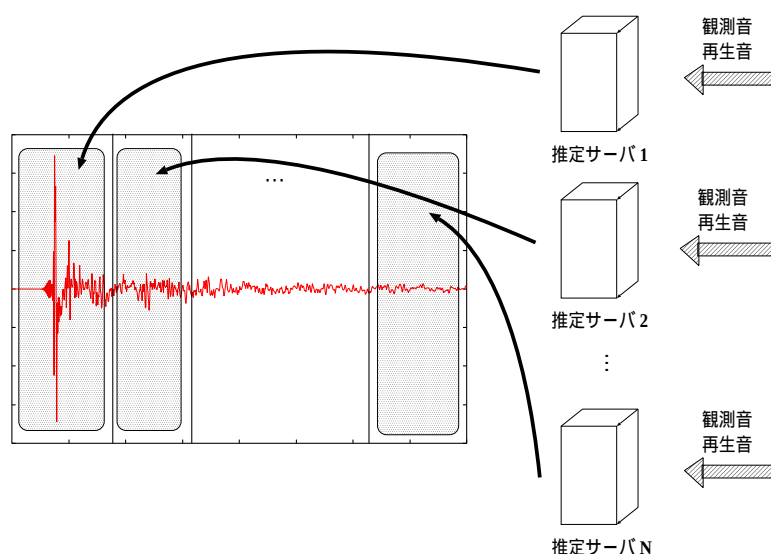


図 4.4 負荷分散手法

4.3.1 実験条件

実験内容としては、補正と推定処理により再現空間の伝達特性を打ち消すことが可能なタップ数を実時間処理可能な限界を確認することで本システムの有効性を示す。使用した機器は図 4.1 となっている。この実験では、これらの機器を用いて本研究で提案する分散処理において用いた推定サーバの台数が最も効率良くなる条件を明らかにする。また各機器は実験環境間を Ethernet 接続されている。

高性能な CPU を搭載している EndeavorPro4000(Core2Duo) と、WS8000GT(Opteron) はデュアルコアとなっているため各々のコアに補正処理のプロセスを割り当てることが可能である。このことにより条件 4 では二台の機器の四個のプロセッサによる分散処理の性能を確認する。また、プロセッサの処理性能は Core2Duo の方がより高性能であり、Opteron との分散処理によって低性能な機器との分散処理が可能かを条件 3 を行い条件 4 と比較することで確認を行う。また、条件 1 と 2 では分散処理を用いなかった場合の処理性能を確かめる。

- 条件 1 クライアントが単体で補正と推定処理を行う。
- 条件 2 ネットワーク上の EndeavorPro4000 で補正と推定処理を行う。

4.3 実験による評価

表 4.1 使用した機器

補正サーバ	EndevorPro4000
CPU	Core2DuoE6700(2.66GHz)
Memory	PC2 5300(2GB)
推定サーバ 1	EndevorPro4000
CPU	Core2DuoE6700(2.66GHz)
Memory	PC2 5300(2GB)
推定サーバ 2	WS8000GT
CPU	Opteron165(1.8GHz)
Memory	PC3200(2GB)
クライアント	-
CPU	Pentium4(3GHz)
Memory	PC2700(1GB)

- 条件 3 補正サーバと推定サーバ 1 により分散処理を行う．
- 条件 4 補正サーバと推定サーバ 1,2 によって分散処理を行う．

4.3.2 実験結果と考察

サンプリングレート $48.000Hz$ の白色信号を入力信号として，各条件で実時間処理可能なタップ数を 4.2 に示す．

クライアントが単体で動作している条件 1 ではタップ数は 1000 となった．この結果と比べネットワーク上の機器により処理を行っている条件 2,3,4 はより処理できるタップ数が向上していることが確認できた．しかし，2 台の推定サーバを使用している条件 4 は条件 3 と比べてタップ数が向上していない．これは，推定サーバ 2 の CPU 処理性能がボトルネックとなっていることが考えられる．考えられる原因として，本システムで採用しているプロッ

4.3 実験による評価

表 4.2 実時間処理可能なタップ数

	タップ数
条件 1	1,000
条件 2	2,000
条件 3	6,000
条件 4	3,000

ク単位のデータに対応した逆システムの推定法は、ブロックデータ毎に 1 度だけフィルタの更新が可能なることから処理に多くの演算量を必要としていない。そのため、推定処理を行うフィルタを分割して各機器に割り当てたとしても負荷分散の効果が少ないことが考えられる。よって、本手法によって構成した実時間音場再生システムは同性能な補正・推定サーバを用いた場合に最も実時間処理性能の向上が可能である。

4.3.3 提案システムの問題点

本研究で提案したシステムでは適応信号処理による原音場の再現は不可能だった。原因としてネットワークやサウンドデバイスによる遅延などが考えられる。この遅延を 1 サンプル毎に 1 ずつ値を増やし、fragsize に達した時に 0 に戻した値を送信したところ図 4.5 のようなデータを受信した。送信データに対応する受信データを見ると、不必要なデータが入ることにより遅延が生じている。適応信号処理は因果律を満たす必要があるが全くランダムなタイミングで遅延が起こる環境では遅延器による所望信号と観測信号の同期は難しい。本手法では TCP によるデータの送受信を行ったが正確な伝送ができなかったことで音場再生が実現できなかったと考える。

4.4 WAN 上での問題点と解決案

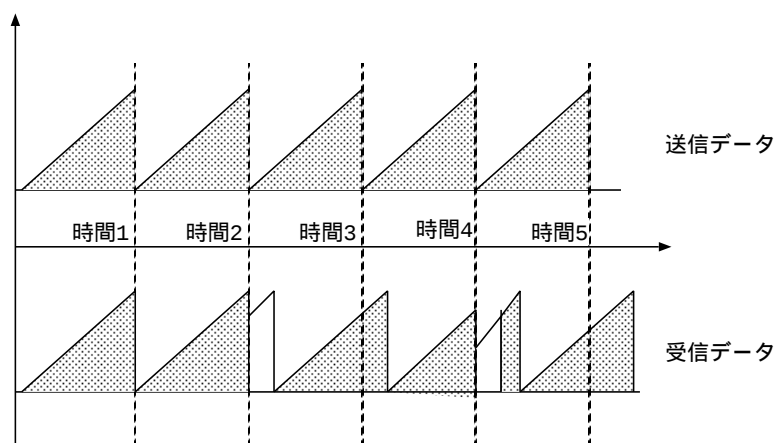


図 4.5 遅延の確認

4.4 WAN 上での問題点と解決案

前節 4.3.2 ではシステムの有効性の確認するため各機器を LAN 内で Ethernet 接続して実験を行っている，結果として補正，推定サーバは同様の処理性能を持つときに最も性能向上が見込まれることが確認することができた．しかし，実際を使用する場合を考えると実験環境のような機器間が Ethernet で接続されており帯域幅が広く遅延の少ない理想的な環境であるとは限らない．そこで本節では一般のネットワークでの利用を考え，狭い帯域幅や通信トラフィックの影響が考えられる環境で本システムを使用した場合について考察を行う．

本システムでは 4.2.3 節で述べたように，補正サーバが再生信号，観測信号そしてフィルタ係数の送受信を行う起点となっている．そのため補正サーバは他の機器に比べてネットワークトラフィックが増加しやすい状況にあるといえる．そこで補正サーバとクライアント間，補正サーバと推定サーバ間でデータの欠損があった場合にシステムがどのような挙動を示し，実時間音場再生の実行に与える影響を考える．そして最後に一般のネットワークで本システムを使用するために必要な機能を考察する．

4.4.1 データの欠落による影響

本システムは TCP でデータの送受信を行っている．TCP はコネクション指向で，信頼性のあるバイト・ストリーム転送が可能である．この信頼性を確保するために TCP はチェッ

4.4 WAN 上での問題点と解決案

クサムによるデータ内容の保証，タイマーによる再転送，通信負荷による輻輳の回避などの機能を持っている．ここで，補正サーバとクライアント間，補正サーバと推定サーバ間を WAN で接続したときを考える．初めに，補正サーバとクライアント間が WAN 接続されている図 4.6 のような構成となっている場合を考える．

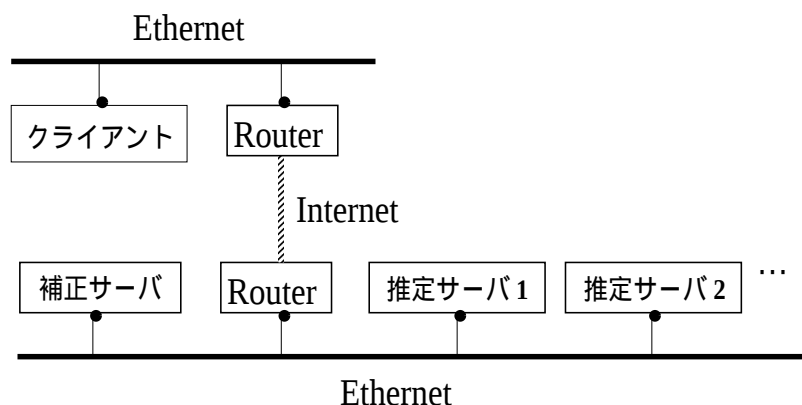


図 4.6 補正サーバとクライアントの WAN 接続

このとき，補正サーバとクライアント間ではイレギュラーな通信負荷や低速な回線があることで再生・観測信号の送受信に影響がでることがある．高速な回線から低速な回線にデータが向かうときや，複数の入力があるルータにあったとき輻輳状態になりやすい．このときルータのキューが受けきれなかったデータは破棄される．データを正確に送信するためには破棄されたデータを再送する必要がある．そのため TCP では再送処理により輻輳状態を悪化させないために再送間隔を長くしパケットの転送率を減らすことで輻輳を回避する機能を持っている．つまり WAN 上のルータと通信状況により補正サーバとクライアント間の遅延量が変化することがわかる．本システムでは処理が逐次的に行っていることから遅延量の増加は実時間処理性能に影響がでる．

次に，補正サーバと推定サーバ間が WAN 接続されているときは図 4.7 のようになる．この場合でも同様に WAN 上の通信状況によって遅延の影響を受ける．しかし補正サーバ・クライアント間が WAN 接続されている場合と異なることは，前者は再生が途切れないよう常に再生信号を送り続ける必要がある．しかし，後者は補正サーバが必要なデータを受け取り推定処理を行っており，通信の遅延により影響が出たとしてもフィルタ係数を受け取る更新タイミングまでに推定処理が終っていればフィルタ係数の更新が可能である．

4.4 WAN 上での問題点と解決案

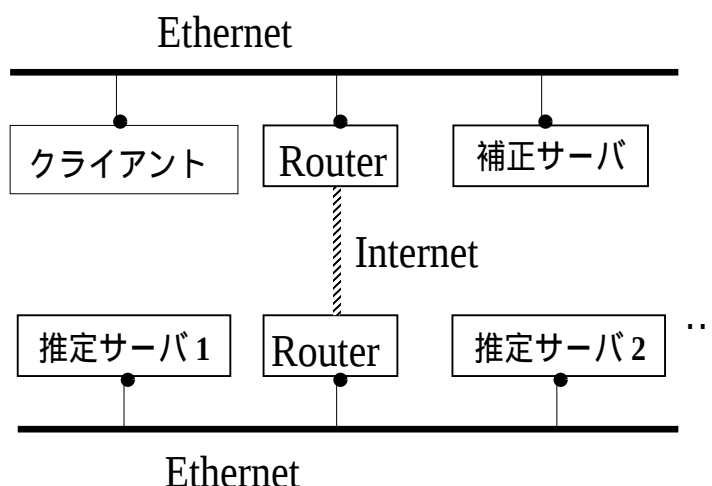


図 4.7 補正サーバと補正サーバの WAN 接続

4.4.2 ネットワーク遅延の解決案

以上述べたように，本システムを WAN 上の機器を用いた動作させた場合，実験環境と比べてネットワークによる遅延を考慮する必要がある．実時間処理を行うにはブロックデータ当たりの再生時間 T を，補正処理に必要な時間 P と送受信に必要なネットワーク伝送時間 N により

$$T \geq P + N \quad (4.1)$$

を満たす必要がある．このとき， N を計測する方法に TCP のタイムアウトと再転送時間を決定する基本となっているコネクションにおける往復時間 (RTT) を用いることで

$$T \geq P + RTT \quad (4.2)$$

と表すことができる．この RTT は経路の変化やネットワークトラフィック量の変化によって時間とともに変化をする．そのため RTT の変化に応じて処理時間 P を増減させる手法が必要となる．考えられる手法として補正を行うタップ数を RTT によって減らすことがある．しかし，タップ数を減らすことで逆特性によって補正できる再現環境の伝達特性が短くなってしまう．またこの手法を取ることで適応信号処理の最適化ができなくなることも考えられる．

4.5 まとめ

本章では，本研究で提案している利用者側の処理負荷低減による実時間音場再生システムの構築手法をサウンドデバイス，音場再生システム，分散処理手法について述べた．また，実験によって本システムの有効性を確認し，結果として2台の高性能な機器による分散処理が最も効率良く処理を行えることがわかった．しかしながら，データの送受信に遅延が生じたことにより音場再生が実現できなかった．また実験環境以外で本システムを実行したときネットワーク上のトラフィックの変化などにより実時間処理性能に影響がでることから，WAN 接続する機器間を変更し各条件で起こり得る影響を考察し，その解決法を示した．

第 5 章

結論

5.1 本研究のまとめ

再現空間において別の空間の音場を完全に再現するには、再現空間における減衰や反響といった影響を除去する必要がある。従来手法では、再生と補正を実時間で実行するには高い処理能力を持つ機器を必要としており容易に実現することができなかった。この問題を解決するため、本研究では利用者の機器に依存しないサービスの一つとして実時間処理可能な音場再現システムを構築した。また実験の結果から、遅延などの影響により音場再生処理は実現できなかったものの、利用者は提案システムを使用することで実時間に処理可能なタップ数が増加していることが確認できた。

5.2 今後の課題

本研究では 2 台のサーバの 4 つのプロセッサを用いて分散処理を行った。結果として補正サーバと同性能な推定サーバを用いることが最も実時間処理能力が高いことが確認でき、低性能な機器の有効活用が十分にできていないことがわかった。このことから今後の課題としてより低性能な処理装置を活用できる分散処理方式を実装する必要がある。またシステム構築には実験環境のような理想的な環境以外でも実現できるようネットワークのトラフィックや転送速度なども視野にいれる必要がある。また、収束特性の観点からブロック適応アルゴリズムの導入による収束特性の改善など、システムの改良すべき点についてさらなる検討が必要である。

謝辞

本研究を行うにあたり，日本語すら危うい私を見捨てず日々夜遅くまで研究室にいて厳しくもあたたかく御指導してくださいました高知工科大学情報システム工学科の福本昌弘准教授に深く心より感謝します．また，本論文を審議してくださいました島村和典教授，浜村 准教授に感謝します．

そして，同大学院生の佐伯 幸郎氏，福富 英次氏，劉 立剛氏，に心から感謝します．佐伯氏には研究はもちろん，研究室での様々なサプライズや，それに付随する料理の数々などお世話になりっぱなしでした．福富氏には，忙しいにも関わらず梗概提出や卒論発表前には夜遅くまで面倒を見て下さりありがとうございました．劉氏には，日頃の癒しと，寮で中国料理を振舞って頂きました．もう少し呑まれないようになりたいところです．

また，一緒に研究活動を戦った福本研究室の一同にも感謝致します．金井君にはちょっとした現実逃避やドライビングなど多くの時間を過ごしました．同期の中一人院への進学となりましたが，この一年間に多くのことを学び天下を築いてくれると信じています．また，野村君には打たれてもくじけない強さを武器に今後の活躍を期待しています．また，小林君には論文の訂正に続く訂正で思考が止まった私を手助けしてくれて助かりました．今後はシャリ製造機の製作に尽力し世界の名工になることを期待しています．

最後になりましたが，これまで支えて下さった全ての方々に感謝致します．

参考文献

- [1] 浜崎信二, 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, vol.104, no.559, pp.31-36, Jan.2005.
- [2] 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
- [3] 山崎芳男 金田豊, 音・音場のデジタル処理, コロナ社, 1999
- [4] 北脇信彦, デジタル通信・オーディオ技術, 電器通信協会, 1999
- [5] 藤井佑一, "DSP による多入力信号補正システムの構築法", 平成 15 年度 学士学位論文, 2003.
- [6] 伊勢史郎, "キルヒホッフーヘルムホルツ積分方程式と逆システム理論に基づく音場制御の原理", 日本音響学会誌 53, pp706-713(1997).
- [7] W・リチャード・スティーヴンス, 詳解 TCP/IP Vol.1, 株式会社ピアソン・エデュケーション, 2000
- [8] ゲリー・R・ライト, W・リチャード・スティーヴンス, 詳解 TCP/IP Vol.2, 株式会社ピアソン・エデュケーション, 2002
- [9] 4Front Technologies, Open Sound System Programmer's Guide version1.11, 2000

付録 A

広い領域での音場再生法

広い領域での音場再現手法として、任意の位置に配置された複数の音源により、ある空間の境界面上の音圧と音圧勾配を制御する境界音場制御がある??。この手法は図 A.1 のように、スピーカによって制御領域を取り囲むように配置し、制御領域の境界面上のマイクで観測した音圧と音圧勾配を等しくすることで再現音場を原音場と等しくすることが可能である。

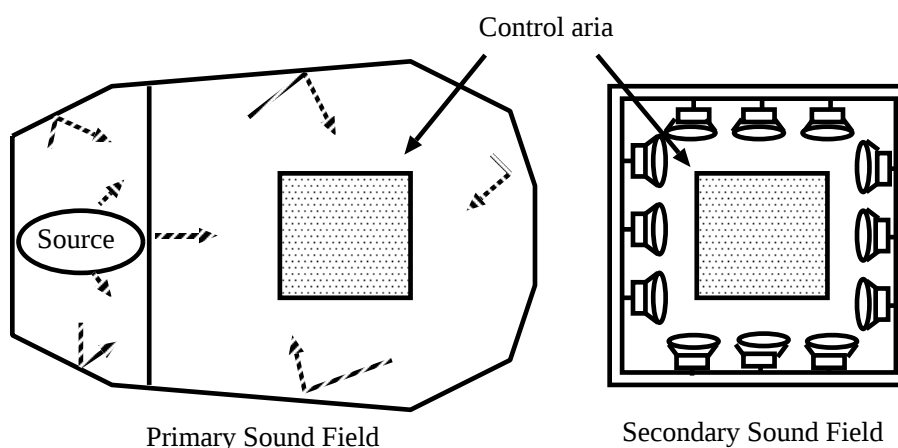


図 A.1 境界音制御の概念図

このとき、スピーカから制御領域の境界面までの伝達特性を打ち消すためにスピーカ数 M 境界面上の制御点 N に対応する $M \times N$ 行列の逆フィルタが必要である。そして、逆フィルタによって伝達特性を打ち消し、制御領域の音圧と音圧勾配を原音場と等しくすることで制御領域内であれば再現音場で原音場を再現することができる。