

平成 20 年度
学士学位論文

ネットワークを利用した音場再生における 分散手法

Dispersion method for sound feild reproduction
system on the network

1090349 徳久 翔太

指導教員 福本 昌弘

2009 年 3 月 6 日

高知工科大学 情報システム工学科

要 旨

ネットワークを利用した音場再生における分散手法

徳久 翔太

ディジタル信号処理技術の向上により，高いサンプリングレートやマルチチャネルで録音された音を体験することができるようになった．しかし，それらの技術のみでは私達は十分な臨場感を得ることができない．なぜなら，スピーカーを通して聞こえる音は，常に反響や減衰によって影響を受ける．そこで，臨場感を伝えるための手段として，近年音場再生技術が注目を受けている．

音場再生とは，リスナーの環境による影響を取り除いて，リスナーに原音を提供する技術である．リスナーの環境特性は温度や湿度などいくつかの要因で常に変化しているため，音場再生技術において適応フィルタのパラメータはリアルタイムで調整される必要がある．この実時間でという要求を満たすためには，リスナー側に高性能な処理装置が要求される．このため，リスナーは簡単に音場再生技術を利用することができない．

そこで，ネットワーク上の計算機リソースを用い，リスナー側で処理を行う必要が無い音場再生システムが提案されている．このシステムではクライアント側では，受信音を再生し観測音を送信するだけで，サーバ側で音場再生処理を行う．そのため，構築の際には，補正サーバと補正フィルタを計算する推定サーバをネットワーク上に配置する必要がある．しかし，このシステムはネットワークの状態を考慮していない．そのため，通信速度が十分な速度でない場合に，ネットワークによる遅延の影響を受け，音場再生処理が行われなくなる．そこで，新たな分散手法を提案している．そして，新しい分散手法を用いたシステムと従来の手法のシステムでシミュレーションを行い，提案する分散処理システムの有効性を示す．

キーワード キーワード 音場再生 分散手法 ネットワーク

Abstract

Dispersion method for sound field reproduction system on the network

Syota Tokuhisa

With development of the digital signal processing technology in recent years, we can experience high sampling rate and multi channel recorded sound which is described as having high realistic. However, we can not feel enough realistic sensation by only these technologies because the sound is always damaged by unexpected reverberation and attenuation when we hear through loudspeakers. To provide more realistic sensations, the sound field reproductions system has attracted attention in recent years.

The sound field reproduction is a technique to provide the original sound for the listener to eliminate the effect of listener's field. Parameters of the sound field reproduction system should be adjusted in real time, because the acoustic characteristic of listener's field is always changing in some reason, such as temperature, degree of humidity and position, etc. To satisfy real time this requirement, The sound field reproduction system needs a high performance processor on the listener's side. With this reason, the listener is too hard to use the sound field reproduction technology easily.

In this research, we proposed the sound field reproduction system with use of the computer resource on the network that does not necessary to process on client or listener's side. To construct this system, the correction server and the estimation server which calculate the correction filter are located on the network. And, the client only plays received sound and sends observed signal. the server side processes. but, this

system has not consider transmission speed. So, when the transmission speed isn't enough speed, the system can't provide the sound field reproduction.

In this paper, we have proposed new distribution technique. And, we have compared existing distribution technique and new one by a simulation. Then, we have shown availabilty of this system. And, the sound feild reproduction system is going to became more versatile system.

key words sound field reproduction, dispersion method, network

目次

第 1 章	序論	1
1.1	研究の目的とその背景	1
1.2	本論文の構成	2
第 2 章	音場再生に用いられている技術	3
2.1	はじめに	3
2.2	バイノーラル録音	3
2.3	音響システムと線形モデル	4
2.4	適応フィルタを用いた逆特性の近似	6
2.5	多チャンネル音場再生システムにおける逆特性近似	8
2.6	2 チャンネル-2 点制御	9
2.7	多入力信号補正法	10
2.7.1	逆特性の近似的推定法	10
2.8	適応フィルタの更新	13
2.9	まとめ	13
第 3 章	ネットワークを利用した音場再生システムの概要と問題点	15
3.1	はじめに	15
3.2	提案システムの構成	15
3.2.1	音場再生処理システムにおける信号の処理	16
3.2.2	音場再生処理システムの問題点	17
3.3	まとめ	18
第 4 章	提案する分散手法の導入と評価	19
4.1	はじめに	19

目次

4.2	ブロック単位でのサウンドの制御	19
4.3	提案する分散手法	20
4.4	分散手法を用いてのシュミレーション	20
4.4.1	実験内容	20
4.4.2	実験結果	21
4.4.3	性能比較	22
4.5	まとめ	23
第 5 章	結論	25
5.1	本研究のまとめ	25
5.2	今後の課題	25
	謝辞	26
5.3	謝辞	26
	参考文献	27
付録 A	ラウドスピーカによる音場再生	28

目次

2.1	バイノーラルシステムの例	4
2.2	閉空間における信号伝搬の様子	5
2.3	図 2.2 の等価回路	6
2.4	室内伝達特性の補正	6
2.5	MINT 理論に基づく音場制御系	7
2.6	MINT 理論に基づく音場制御系	8
2.7	2 チャネル-2 点制御系モデル	10
2.8	2 チャネル-2 点制御系モデル (クロストーク無し)	11
2.9	逆近似フィルタと構成図	12
3.1	音場再生システムの構成	16
3.2	音場再生システムにおける処理の流れ	17
4.1	内部バッファ	19
4.2	遅延が発生しない場合	22
4.3	小さい値の遅延がまれに発生する場合	22
4.4	小さい値の遅延が頻繁に発生する場合	23
4.5	大きい値の遅延がまれに発生する場合	23
4.6	大きい値の遅延が頻繁に発生する場合	24
A.1	トランスオーラルシステムの例	28
A.2	MINT 理論によるトランスオーラルシステム	29

表目次

4.1 使用した音源	21
4.2 与えた遅延の値と頻度	21

第 1 章

序論

1.1 研究の目的とその背景

近年、デジタル信号処理技術の向上にともない、高いサンプリングレートやマルチチャネルで記録された音源を再生することにより臨場感のある体験ができるようになった。しかしながら、それらの技術のみでは私達は十分な臨場感を得ることができない。なぜなら、スピーカーを通して耳に聞こえる音は、常に反響や減衰によって影響を受けているからである。そこで、臨場感を高める手段として、音場再生技術が注目を受けている。

音場再生とは、リスナーの環境による影響を取り除いて、リスナーに原音を提供する技術である。その際、リスナーの環境特性は温度や湿度、リスナーの位置などの要因で常に変化しているため、音場再生における補正フィルタのパラメータはリアルタイムで調整される必要がある。しかし、通常の音場再生システムでは、実時間による適応信号処理に高性能な処理装置が必要とされるため、リスナーは容易に音場再生を利用することができない。また、リスナーの端末が複数台ある場合は各端末に高性能な処理装置が必要とされる。そこで、リスナーの端末の性能に関係なく音場再生を提供するシステムが提案されている [1]。しかし、提案されたシステムは、ネットワークやデバイスによる遅延による影響を大きく受けることや、複数のサーバを十分に活用できないなどの問題を

そのため、本論文では、提案された音場再生システムにおいて新たな分散手法を用いることで、元となるシステムの問題点である遅延による影響を軽減し、WAN などのネットワークへの導入を目指すことで、より汎用性の高いシステムとなることを目的としている。

1.2 本論文の構成

論文の構成を以下に述べる．2章では，音場再生処理に用いられる録音方法や音場再生における基礎的な技術である適応信号処理について述べ，基本的な音場再生処理について説明している．また，本研究で用いられている2チャンネル2点制御による音場再生について述べ，適応フィルタの更新について説明している．第3章では，ネットワークを利用した音場再生処理システムの説明を行い，第4章で新たに分散手法の導入，評価を行う．そして最後に，これまでの結果を元に第5章で結論としてまとめとこれからの課題について述べている．

第 2 章

音場再生に用いられている技術

2.1 はじめに

一言に音場再生処理技術といっても様々な技術から成り立っており，実際に音場再生処理を行う音場再生システムは，目的の違いから，システムを構成するハードウェアや再生手法が異なり，様々なシステムが考案されている．本章では，本研究で使用した音場再生システムにおいて基礎的な部分となる録音方法や適応信号処理について説明し，そして，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している．

2.2 バイノーラル録音

原音場を忠実に再現するためには，まず，原音場に忠実な音源が必要となる．原音場に忠実な音源とは，原音場での反射や残響といった影響を受けた音である．しかし，原音場でただステレオ録音するだけでは，そのような音源にはならない．なぜなら，実際に聞こえた音というのは耳に入って来た音であるため，耳の形状による影響も受けることになる．そのため，通常のマイクロフォンで録音するだけでは十分とはいえない．そこで，原音場で聞こえる音を録音する方法としてバイノーラル録音がある．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．また，この録音方法を用いて原音場を再現するシステムにバイノーラルシステム 2.1 などがある．バイノーラルシステムであれば簡単な構成で原音場を再現することができるが，ヘッドフォンによる影響特性を受けてしまうという問題点がある．本研究で用いる音

2.3 音響システムと線形モデル

場再生システムはバイノーラル録音を用いるが、バイノーラルシステムとは異なり、ヘッドフォンではなくラウドスピーカを用いた原音場の再現を目指している。次に、音場再生にお

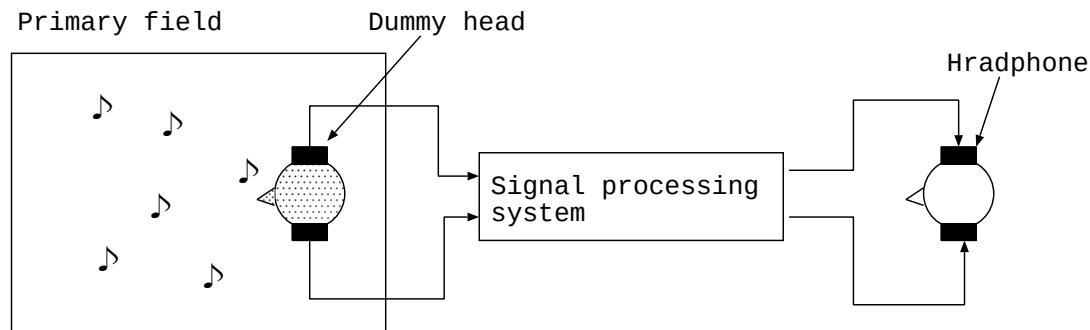


図 2.1 バイノーラルシステムの例

いてもっとも重要な技術の1つである適応信号処理技術について説明する。

2.3 音響システムと線形モデル

音源と受聴点など空間中の2点間の音の伝わり方を示す関数を空間伝達関数 $G(z)$ といい、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、原音場が図 2.2 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信した信号を $y(t)$ 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信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

2.4 適応フィルタを用いた逆特性の近似

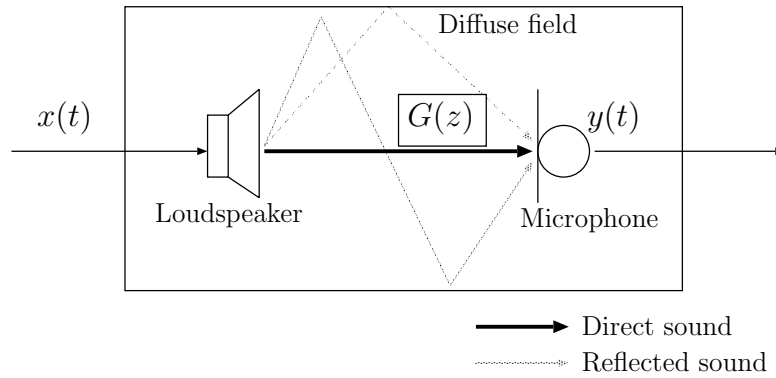


図 2.2 閉空間における信号伝搬の様子

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 2.2 伝達系は図 2.3 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る音響システムは線形モデル化することができる．

2.4 適応フィルタを用いた逆特性の近似

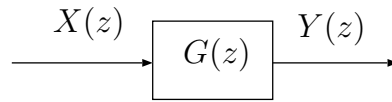


図 2.3 図 2.2 の等価回路

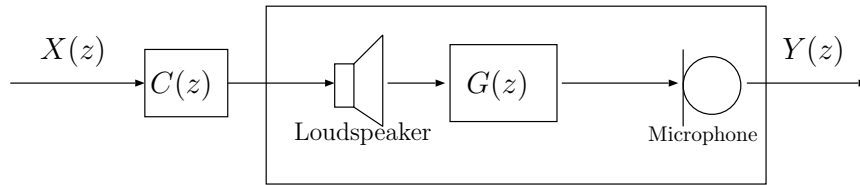


図 2.4 室内伝達特性の補正

2.4 適応フィルタを用いた逆特性の近似

本章の冒頭で説明したバイノーラルシステムと異なり，ラウドスピーカによる原音場の再現を行うシステムのひとつにトランスオーラルシステムがある．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．トランスオーラルシステムにおいては，付録にて説明を行っているので必要であれば参照していただきたい．つまり図 2.3 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，2.4 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 2.4 の関係から補正フィルタと伝達特性を畳み込むことで

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

2.4 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．

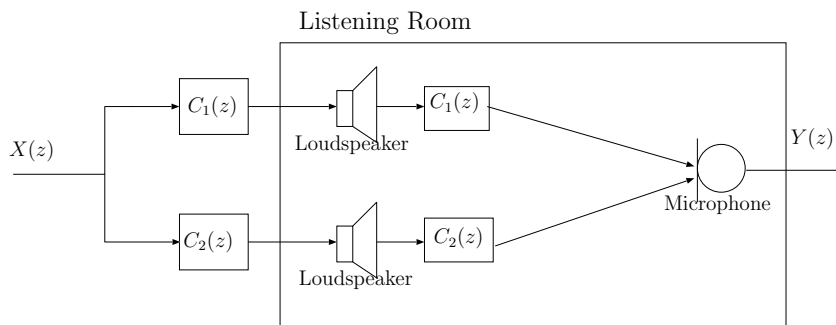


図 2.5 MINT 理論に基づく音場制御系

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 A.2 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

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

2.5 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．

2.5 多チャンネル音場再生システムにおける逆特性近似

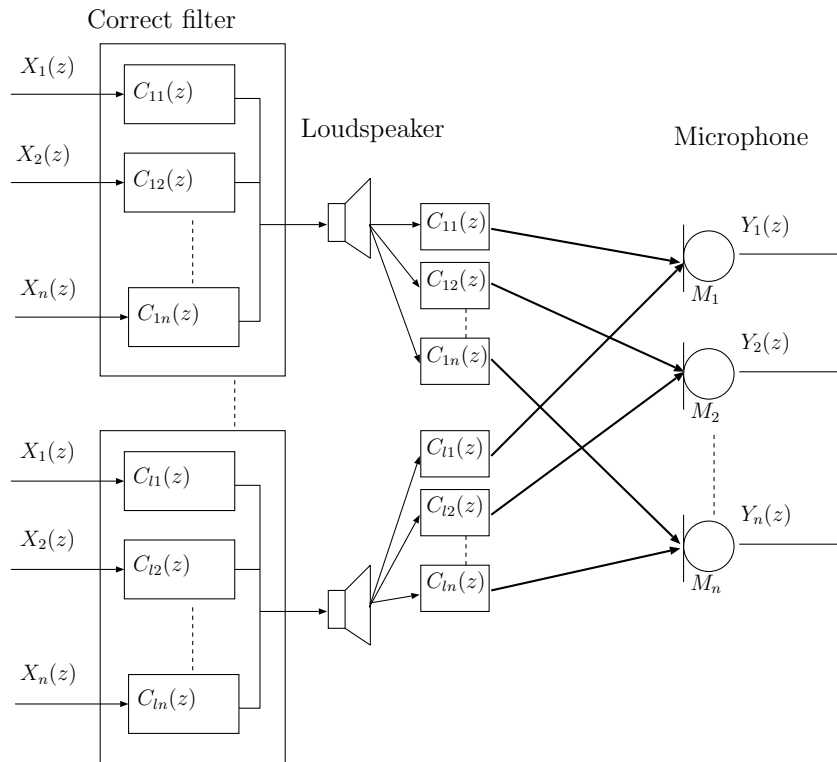


図 2.6 MINT 理論に基づく音場制御系

図 2.6 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

2.6 2 チャネル-2 点制御

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．

2.6 2 チャネル-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樣に等価できないことを示す．図 2.7 に 2 チャネル-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．

$$y_1(t) = x_1(t)g_{11} + x_2(t)g_{21} \quad (2.12)$$

$$y_2(t) = x_2(t)g_{22} + x_1(t)g_{12} \quad (2.13)$$

ここで式 2.13 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

2.7 多入力信号補正法

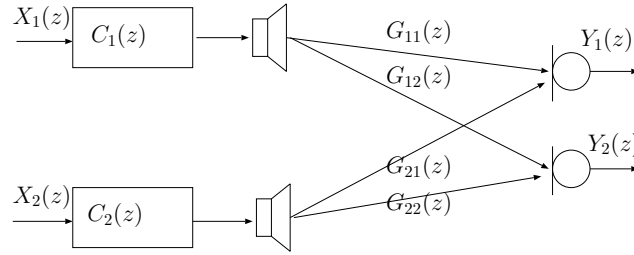


図 2.7 2 チャンネル-2 点制御系モデル

た，式 2.13 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．

2.7 多入力信号補正法

本節ではまず，再生空間においての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再生空間における逆特性を推定する手法について説明をする．

2.7.1 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ていく．説明を簡略化するために，ここでは図 2.8 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．

$$d_i(t) = c_{i,L} x_{i,L}^T(t) \quad (2.14)$$

ここで

$$c_{i,L} = [c_{i,0}, c_{i,1}, \dots, c_{i,L}] \quad (2.15)$$

$$x_{i,L} = [x_i(t), x_i(t-1), \dots, x_i(t-L)] \quad (2.16)$$

2.7 多入力信号補正法

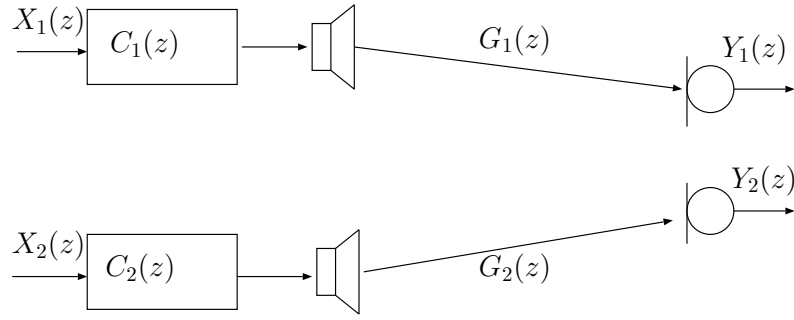


図 2.8 2 チャンネル-2 点制御系モデル (クロストーク無し)

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

$$y_j(t) = d_{i,L_g}(t) g_{ij,L_g}^T \quad (2.17)$$

となる．ただし

$$d_{i,L_g}(t) = [d_i(t), d_i(t-1), \dots, d_i(t-L_g)] \quad (2.18)$$

$$g_{ij,L_g} = [g_{ij}(0), g_{ij}(1), \dots, g_{ij}(L_g)] \quad (2.19)$$

である．この際に補正フィルタ $C(z)$ が伝達関数 $G(z)$ に対する近似逆フィルタと同様の性質を持っていると仮定したならば， 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．

$$\begin{aligned} Y_j(z) &= G_{ij}(z) D_i(z) \\ &= G_{ij}(z) C_i(z) X_i(z) \\ &\approx G_{ij}(z) g_{ij}^{-1}(z) x_i(z) \\ &= x_i(z) \end{aligned} \quad (2.20)$$

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 2.9 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

$$s_{ij,L}(t) = h_{ij,L}^T(t) y_{ij,L}(t) \quad (2.21)$$

2.7 多入力信号補正法

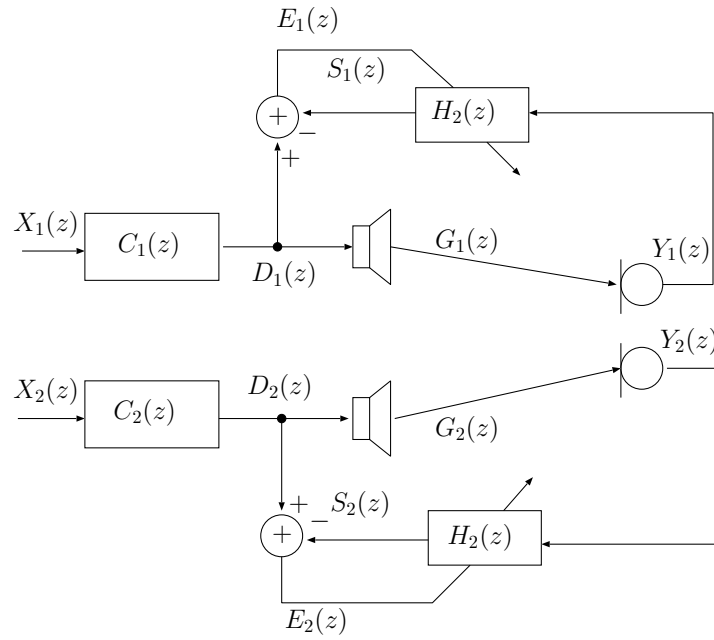


図 2.9 逆近似フィルタと構成図

で与えられる。ただし

$$h_{ij,L}(t) = [h_{ij,0}(t), h_{ij,1}(t), \dots, h_{ij,L}(t)] \quad (2.22)$$

$$y_{j,L}(t) = [y_j(t), y_{t-1}(t-1), \dots, y_j(t-L)] \quad (2.23)$$

である。一方、系への入力信号を式 2.14 で与えられるものとする、出力誤差 $e_{ij}(t)$ は

$$e_{ij}(t) = d_i(t) - s_{ij}(t) \quad (2.24)$$

となる。ここで、出力誤差が $e(t) = 0$ の場合を考えると

$$y_{j,L}(t) h_{ij,L}^T(t) = x_{i,L}(t) c_{j,L}^T(t) \quad (2.25)$$

となる。式 2.25 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、

$$\begin{aligned} y_{j,L}(t) h_{ij,L}^T(t) &= x_{j,L}(t) h_{ij,L}^T(t) \\ y_{j,L}(t) &= x_{i,L}(t) \end{aligned} \quad (2.26)$$

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

2.8 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

$$h_{ij,L}(t+1) = h_{ij,L}(t) + \alpha \frac{y_{j,L}(t)}{\|y_{j,L}(t)\|^2} e_{ij}(k) \quad (2.27)$$

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．

2.8 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．

$$e_{ij}(t) = d_i(t) - s_{ij}(t) \quad (2.28)$$

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．

$$h_{ij,L}(t+1) = h_{ij,L}(t) + \alpha \frac{y_{j,L}(t)}{\|y_{j,L}(t)\|^2} e_{ij}(k) \quad (2.29)$$

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．

$$c_{1,L}(t+1) = h_{11}(t) \quad (2.30)$$

$$c_{2,L}(t+1) = h_{22}(t) \quad (2.31)$$

2.9 まとめ

本章では、音場再生システムに用いられる録音方法であるバイノーラル録音、再生空間における逆特性を推定し補正を行う適応信号処理の基礎について説明した．そして、2 つの音

2.9 まとめ

源を用いて 2 点に対して制御を行う 2 チャンネル 2 点制御を説明し，補正サーバにおける適応フィルタの更新について説明を行った．これらを用いて，音源を録音し，適応信号処理により再生環境の受聴点における逆特性を持った信号に補正することができれば，原音場を再現することが可能であることを説明した．

第 3 章

ネットワークを利用した音場再生システムの概要と問題点

3.1 はじめに

前章までに述べた技術を用いることで、音場再生を実現することが可能である。しかし、序論でも述べたように音場再生処理には、音場再生を行う端末に高性能な処理装置が必要である。そのため、複数の端末で音場再生処理を利用したい場合、各端末に高性能な処理装置が必要となり実用的でない。そこで、ネットワークを利用した音場再生処理システムが提案されている。

本章では、始めに提案されている音場再生システムの全体構成を示している。そして、音場再生システムの実装手法と分散処理の手順を述べ、このシステムにおける問題点を挙げている。

3.2 提案システムの構成

ネットワーク全体の構成を図 3.2 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、再生環境は、クライアントとなる端末、補正音を再生するためのスピーカ、再生音を観測するためのマイクから構成される。クライアント側に求められる最低限の機能として、音を再生しながら受

3.2 提案システムの構成

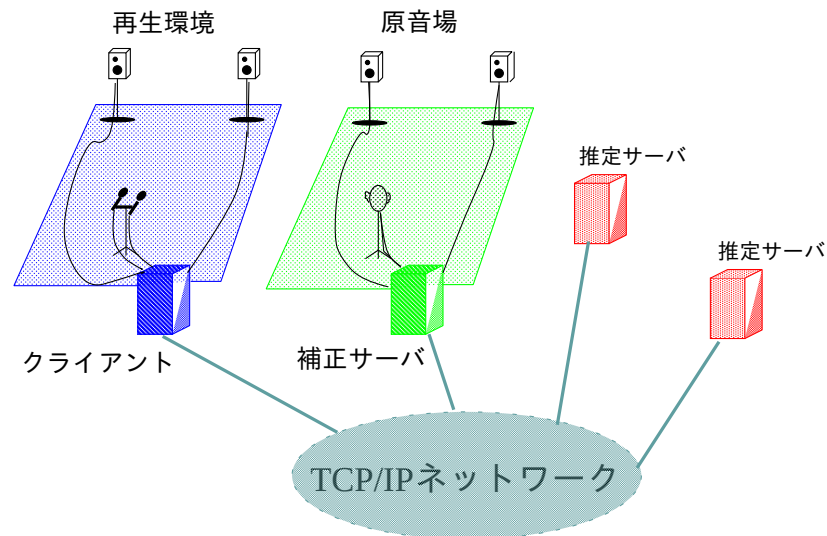


図 3.1 音場再生システムの構成

聴点に設置したマイクで録音を行うことが可能であることが求められる。

原音場となる補正サーバ側は、補正サーバ用の端末と原音を観測するためのマイクで構成される。そして、ネットワーク上に1台以上の推定サーバを構築する。次に、この音場再生システムにおける処理の流れを説明する。

3.2.1 音場再生処理システムにおける信号の処理

ネットワークを利用した音場再生処理システムの処理の流れを図に示す。

- クライアントがサーバに接続を行いコネクションを確立する
- 補正サーバから信号を受け取り、クライアントがスピーカーから再生を行う
- 同時にクライアントは、受聴点のマイクによる録音を行い、観測された信号を補正サーバに送る
- 補正サーバはクライアントから信号を受け取るとその信号をさらに推定サーバに送信する(推定サーバが複数存在するときには信号を分割して送信する)
- 推定サーバはクライアントの環境における逆フィルタのパラメータを推定し、補正サー

3.2 提案システムの構成

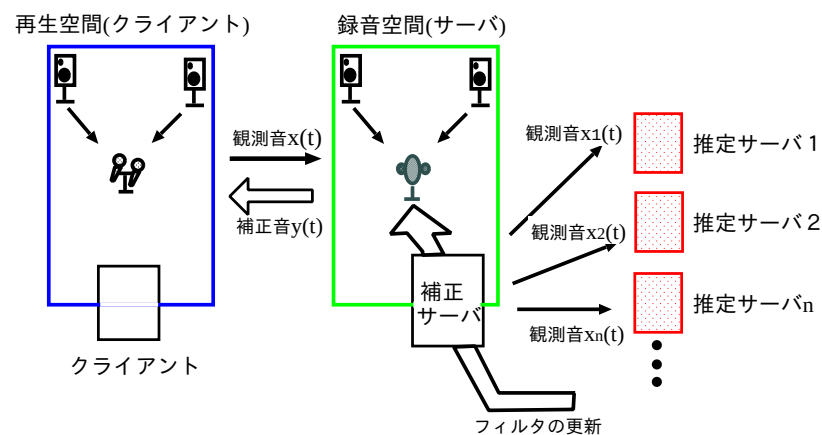


図 3.2 音場再生システムにおける処理の流れ

バに送信する

- 補正サーバは受け取ったパラメータを元に適応フィルタを更新する
- 補正サーバが更新された適応フィルタを用いて、原音に再生空間の逆特性を持たせる
- 処理された信号をクライアントに送信しクライアントが受信した信号を再生を行う

このような流れで音場再生処理を実現している。

3.2.2 音場再生処理システムの問題点

リスナーの環境は常に変化しているので、音場再生処理は実時間で行われるのが好ましい。しかし、このシステムでは、Ethernet 接続を前提としているため、通信トラフィックの影響を考慮していない。その結果、通信トラフィックによる遅延の影響を受けやすく不安定である。遅延の影響によっては、このシステムでは、実時間での適応信号処理が行われなくなってしまうことがわかっている。そのため、WAN での使用には不安が残る、汎用性に乏しいという問題点が挙げられる。

3.3 まとめ

3.3 まとめ

本章では，ネットワークを利用した音場再生処理システムにおけるクライアント側とサーバ側の構成を示し，それぞれが行う処理を説明し，全体の流れを示した．また，システムの持つ問題について述べた．

第 4 章

提案する分散手法の導入と評価

4.1 はじめに

本章では，前章の音場再生システムの分散手法を新たに提案する．そして，提案手法を用いることでより汎用性の高いシステムとなることを，以前の分散手法を用いたシステムとの計算機シミュレーションの結果を比較することで示している．

4.2 ブロック単位でのサウンドの制御

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．また，このデバイスを OSS(Open Sound System) によって制御する．

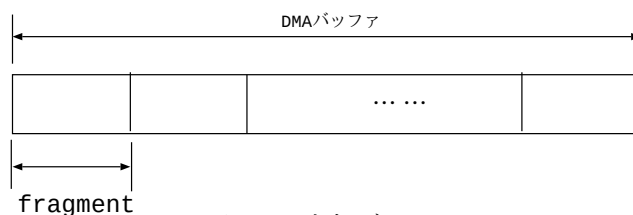


図 4.1 内部バッファ

サウンドカードは図 4.1 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(Open Sound System) 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．

- 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．

4.3 提案する分散手法

よって、本システムでは fragsize のブロックサイズ毎に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ことで実時間性を維持している。

4.3 提案する分散手法

リスナーの環境は常に変化しているので、音場再生処理は実時間で行われるのが好ましい。しかし、変更前の分散手法の問題点は、Ethernet 接続を前提としているため、狭い帯域幅や通信トラフィックの影響を考慮していない。そこで、WAN 接続にも利用できるような TCP のコネクションにおける往復時間 T_{RTT} を用いて分散処理を行う。実時間実行処理を行うにはブロックデータ当たりの再生時間を T_{Play} 、補正処理に必要な時間を T_{Pro} とすると、

$$T_{Play} \geq T_{Pro} + T_{RTT} \quad (4.1)$$

を満たすように処理を行う必要がある。よって、 T_{RTT} の変化に応じて、補正を行うタップ長を調整し T_{Pro} を変化させる。この手法を用いることで、ネットワークの通信トラフィックを考慮した分散処理を実現した。

4.4 分散手法を用いてのシミュレーション

提案した分散手法を用いた音場再生システムと分散手法を変更しない音場再生システムを想定した計算機シミュレーションを行った。

4.4.1 実験内容

ネットワークの遅延が存在する場合、それが音場再生処理にどれほど影響するか計算機によるシミュレーションを行った。実験内容は、補正サーバとクライアントをそれぞれ実行し、補正時に遅延を故意に発生させる。この際、式 (4.1) を満たすように処理を行うブロック数を変化させたものとそのまま処理を行った時の原音と観測音との差を求めることで比較

4.4 分散手法を用いてのシミュレーション

を行った。

実験に使用した音源の詳細については表 4.1 に示している。

表 4.1 使用した音源

量子化ビット	64bit
サンプリングレート	44.1kHz
チャンネル数	1
原音	音声信号
外乱	白色雑音

表 4.2 与えた遅延の値と頻度

———	遅延のサイズ	発生頻度 ‘
シミュレーション 1	発生しない	発生しない
シミュレーション 2	10bit	200bit 毎程度
シミュレーション 3	10bit	50bit 毎程度
シミュレーション 4	44100	10000bit 毎程度
シミュレーション 5	44100	10000bit 毎程

4.4.2 実験結果

遅延の度合を変更してシミュレーションを行い、横軸には bit 番号、縦軸には原信号と観測信号の正規化ノルムをとして、グラフで表している。実線で書かれたグラフが提案手法を用いたもので点線のグラフは従来の分散手法を用いたものである。各グラフは図 4.2 の条件で行っており、図 4.6 がシミュレーション 1 の条件、図 4.5 がシミュレーション 2 の条件以下シミュレーションの番号と図番号が対応している。

4.4 分散手法を用いてのシミュレーション

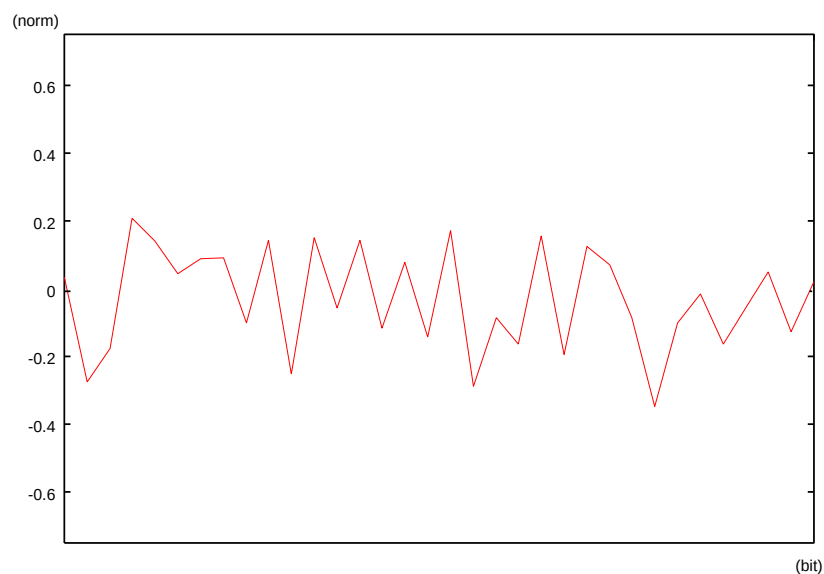


図 4.2 遅延が発生しない場合

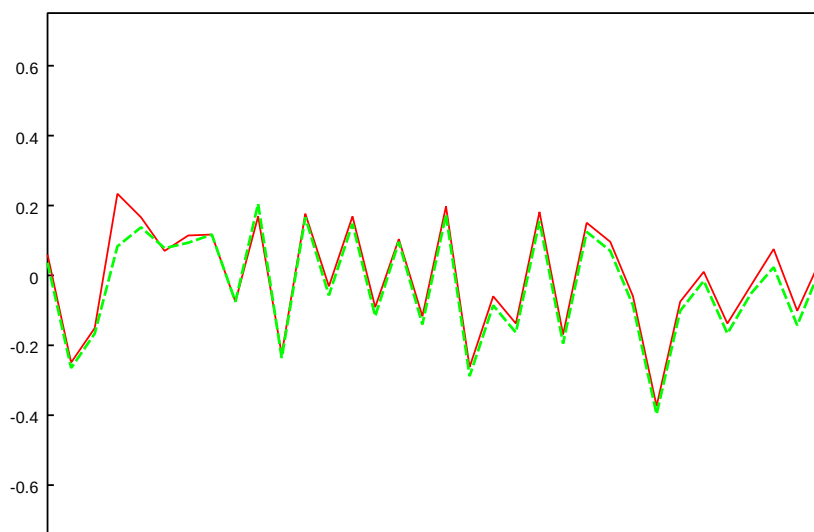


図 4.3 小さい値の遅延がまれに発生する場合

4.4.3 性能比較

従来手法でも提案した手法でも小さい値の遅延の時はそれほど収束特性に変化はないが、遅延の値が大きくなった場合に従来の分散方法はかなり不安定になってしまう。一方、ブロックサイズを変更する提案手法を用いられた場合、影響は受けるが従来手法よりも安定して音場再生処理を行っていることがわかる。

4.5 まとめ

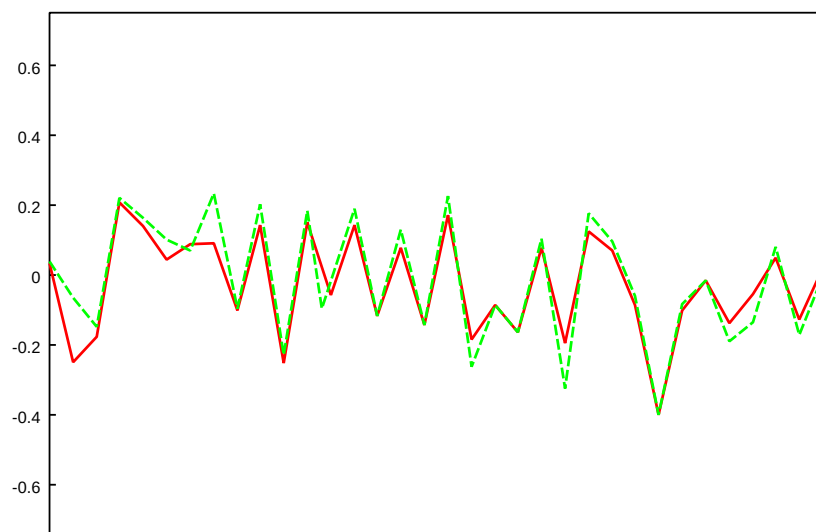


図 4.4 小さい値の遅延が頻繁に発生する場合

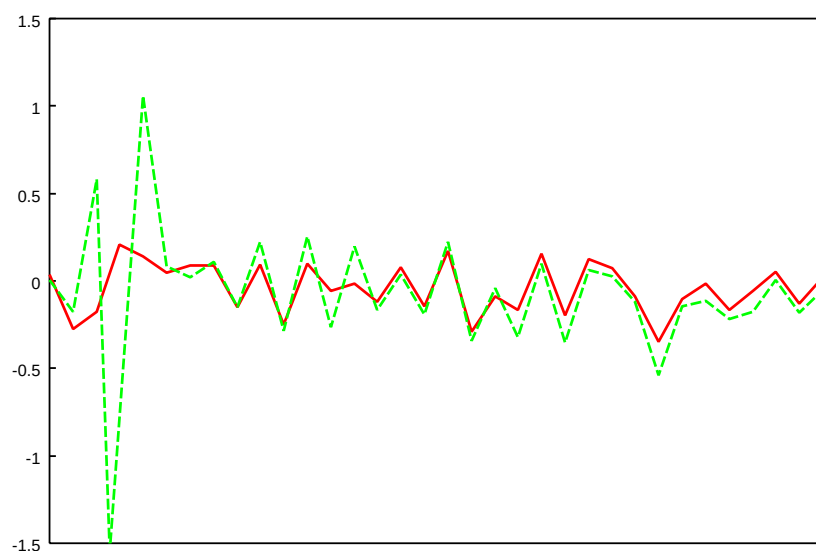


図 4.5 大きい値の遅延がまれに発生する場合

4.5 まとめ

本章では，新たにパケットの往復時間によって補正を行う信号を変化させる分散手法を提案，導入し，計算機シミュレーションを行った．実験結果より，遅延により音場再生システムは少なからず影響を受ける．その際，従来の分散手法を用いるとネットワークの速度が遅い場合に発生する遅延の影響を受けやすく，不安定である．一方，提案した分散手法を用い

4.5 まとめ

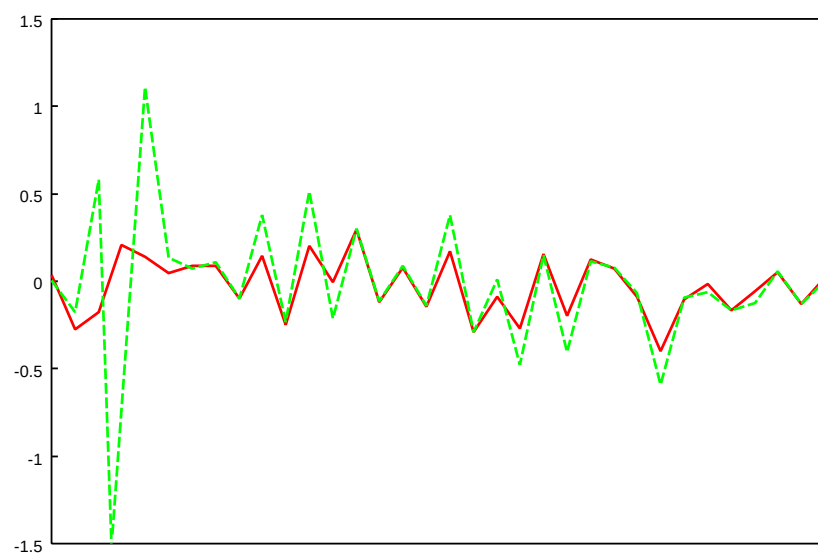


図 4.6 大きい値の遅延が頻繁に発生する場合

ることで従来手法より安定しており，従来の分散手法よりもネットワークによる遅延を軽減することができたと考えられる．

第 5 章

結論

5.1 本研究のまとめ

実際のコンテンツとしてクライアントに対して音場再生を提供する場合，ネットワークにおける速度は一定でない場合がほとんどであると考えられる．しかし，従来手法の分散方法では，ネットワークの速度の変化を考慮しておらず一定のブロックサイズで信号を送信していた．そこで，新たにソケットの往復時間を考慮してブロックサイズを変化させる新たな分散手法を提案した．この分散手法を用いることで，ネットワークが一定の速度でなくてもより安定して音場再生処理を行えることがシミュレーションの結果からわかった．

5.2 今後の課題

現在では，まだ計算機によるシミュレーションの結果でしかないので，実際の環境で使用した場合にどのような結果となるのか観測する必要がある．そして，実際の環境で構築し修正を行うことでより，安定した音場再生システムへと改良していく必要がある．

また，実際の通信を利用して，通信速度とブロックサイズを変化させる実験を行い，一定の速度毎に処理可能なタップ長が増加するブロックサイズを調べる．その結果を利用し，通信速度が十分な場合でもより，効率的な音場再生処理を行うことが可能であると考えられる．

謝辞

5.3 謝辞

本研究を行うにあたり，研究室で指導にあたって頂いた福本昌弘准教授に厚く御礼申し上げます。進学後するからにはしっかりと指導を受け，少しでも多くの能力を身に付け卒業時には惜しまれるような人材になれることを目指してまずはメ切を守ることから頑張ります。そして，卒業研究において，副査をしていただいた浜村昌則准教授，吉田真一講師，学内だけでなく学外においても御指導をいただいた妻鳥貴彦講師に厚く御礼申し上げます。

福本研究室におきましては，院生の方々に大変お世話になりました。ここに御礼申し上げます。博士後期課程の佐伯幸郎氏には日々の研究以外に多くのことを教えて頂きました。佐伯氏の行動力にはよく驚かされましたが，研究への活力となりました。同福富英次氏には著者が大学3年時から輪講等で懇切丁寧に御指導頂きました。修士1年の金井宏一郎氏には研究だけにとどまらず様々な御指導をいただきました。金井氏の先生に対しても物怖じせず接することのできる度胸には憧れを抱きます.. また，本当にいろいろありましたが，研究室とともに研究を行った同期の青野君，岡崎さん，越知君に御礼申し上げます。

福本研究室以外においては，著者が卒業研究を行うにあたり間食気味になり，体型に関して気になった時には，吉田研究室の方々に励ましていただきました。そして，気分によって急に様々な所に誘う筆者につき合ってくれた友人達に厚く御礼申し上げます。一足先に社会にでる友人たちの活躍を心から願っています。

最後に，いままで私をささえてくれたすべてに感謝致します。

参考文献

- [1] 山田良, 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, 高知工科大学, 2007
- [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] 辻井重男, 鎌田一雄, デジタル信号処理, 昭晃堂, 1990
- [4] 山崎芳男 金田豊, 音・音場のデジタル処理, コロナ社, 1999
- [5] 北脇信彦, デジタル通信・オーディオ技術, 電器通信協会, 1999
- [6] 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.
- [7] Jeff Tranter, Linux マルチメディアガイド, オライリージャパン, 1997.

付録 A

ラウドスピーカによる音場再生

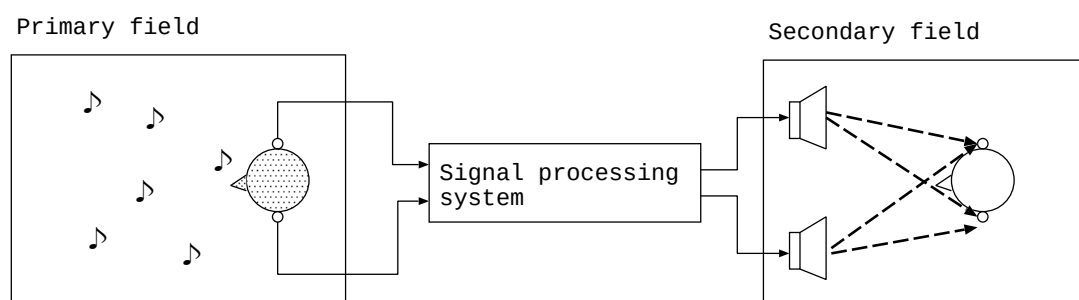


図 A.1 トランスオーラルシステムの例

バイノーラル収録された原信号を，ラウ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ステム図 A.1 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．

しかし，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するコスト，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

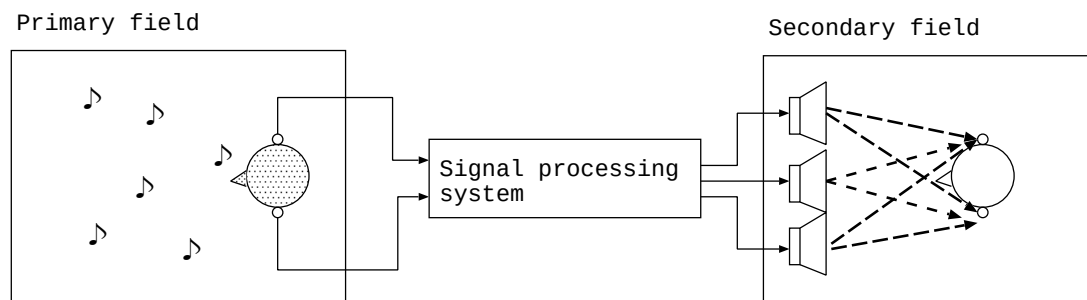


図 A.2 MINT 理論によるトランスオーラルシステム

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．