

# 博士論文

多点制御法を用いた音のエリア再生に関する研究

2020年1月

龍谷大学大学院 理工学研究科

博士後期課程 情報メディア学専攻

T16D001 安枝 和哉



# 多点制御法を用いた音のエリア再生に関する研究

安枝 和哉

## 内容梗概

近年、計算機の性能向上により、高解像度な映像や臨場感のある音響を処理できるようになり、VR(Virtual Reality)やAR(Augmented Reality)などで多くのコンテンツが作成されている。その際、音の方向や大きさを制御することは臨場感を与える重要な要素となる。現状、VRなどではヘッドホンやイヤホンなどを用いて個別に音を再生することが主流であるが、これらを使用せずに音を制御することができれば、より高臨場感を与えられることが期待できる。そこで、複数のスピーカを用いて音場を制御する音場制御の研究が行われている。これらの研究は、収録した音場を別の領域で再現することを目的としているが、この理論を応用することで音を特定の方向や領域のみに再生する、音のエリア再生を行うことができる。

音のエリア再生は、特定の情報を必要な人のみにピンポイントで届けることができるため、駅ホームでの案内放送や美術館・博物館などでのガイダンスシステムなどに利用することができる。また、複数の場所に異なる音声を再生することで、多言語による案内放送やデジタルサイネージへの応用も期待できる。本論文では、複数のスピーカを用いて領域内の音圧を制御することで領域内の任意のエリアへの再生と複数領域への再生を目的とし、多点制御法を用いた音のエリア再生について、その再生手法を論じる。

本論文は7章から構成される。第1章では、研究背景と目的について述べ、課題と目的に対するアプローチを説明する。また、単一のスピーカユニットを用いる方法と複数のスピーカを用いる音の指向性制御の方法について説明を行い、それぞれの課題について述べる。そして、本研究で用いる音場制御手法について説明し、本論文の構成

を述べる。

第2章では、複数のスピーカを用いて音場を制御する手法とその原理を応用したエリア再生の手法について述べる。ある領域において、任意の点の音圧はキルヒホッフ-ヘルムホルツ積分方程式より領域境界の音圧と音圧勾配によって求めることができることが知られている。この理論から領域外に設置した複数の音源を用いて、領域内の音圧を制御する境界音場制御が提案されている。境界音場制御では、閉じた領域において音圧か音圧勾配のどちらか一方の制御によって、もう一方も同時に制御できることを示している。また、領域外に設置した音源を用いて、領域内の音圧を所望の音場と一致させる多点制御法がある。本研究では、多点制御法を用いて、空間内に音圧が大きい場所と小さい場所を生成する。領域内における制御を行う点を制御点と呼び、再生を行う制御点を再生制御点、抑圧を行う制御点を抑圧制御点とする。本章では、多点制御法によるエリア再生手法について、相対音圧レベルによる評価と主観評価実験を行い、再生を行うエリア外において音声内容が聞き取れる可能性があることを示した。

第3章では、音声聞き取れる範囲を狭くするため、信号分解法と多点制御法を組み合わせ合わせたエリア再生手法を提案する。入力信号を聞き取れないサブ信号に分解し、それぞれを多点制御法によって同じ場所に再生を行う。その場所ではサブ信号が再合成されて聞き取れるようになり、それ以外の場所ではサブ信号の位相がずれることで、聞き取りを困難にすることができる。また、多点制御法を組み合わせることで、目的のエリア外で音圧を低下させ、不快感の低減を検討する。本章では、サブ信号生成の従来手法としてランダムベクトル分解法を説明する。また、入力信号をLPC(Linear Prediction Coding)分析によってLPC係数とLPC残差信号に分解し、残差信号を用いて複数のサブ信号を生成する手法を提案する。相対音圧レベルによる評価と主観評価実験によって提案手法の有効性を確認した。

第4章では、多点制御法において指向性が十分に形成できず、周囲に漏えいする音声に対して、マスキングを行う手法を提案する。多点制御法を用いたエリア再生手法において、再生を行うエリア外で音声内容が聞き取れる問題があった。この問題に対して、再生対象外の領域にマスキング音を再生することで音声内容が聞き取れないよ

うにする。しかし、再生したマスキング音が再生対象のエリアに漏えいし、本来の再生エリアにおいて聞き取り難くなる可能性がある。そこで、マスキング音の再生についても多点制御法を用いて制御を行い、対象エリアへの漏えいを低減する手法を検討した。主観評価実験によって提案法の有効性を確認した。

第5章では、多点制御法を用いて、1つのスピーカアレーから異なる領域に異なる音声を再生するための制御点配置について提案する。多点制御法は領域内に設置した制御点での音圧を制御することで指向性を形成する。そのため、制御点数が多くなると伝達関数の計測コストや逆行列演算による誤差が大きくなり、音声品質が低下する可能性がある。したがって、より少ない制御点数で広い領域を制御できることが望ましい。そこで、制御点を制御領域の境界とスピーカアレーに近い位置のみにL型に配置することを検討した。また、実際にスピーカアレーを作成し、実環境における実験を行った。実測した伝達関数を用いた場合と自由空間の伝達関数を用いた場合について相対音圧レベルによる比較した。さらに、単語理解度試験を行い、自由空間伝達関数を用いた場合であっても十分な性能が得られる可能性を示した。

第6章では、遺伝的アルゴリズムを用いて制御点配置を決定する手法を提案する。第5章で述べた複数エリア再生について、より効率的に制御を行うため、遺伝的アルゴリズムによる配置の決定方法について検討を行った。また、遺伝的アルゴリズムの評価関数に、客観的な音質評価指標である PESQ(Perceptual Evaluation of Speech Quality)を導入し、再生点付近での音質を考慮した制御点配置についても検討を行った。計算機シミュレーションによる評価実験を行い、エリア再生の性能の向上を確認した。また、得られた制御点配置が第5章で提案した制御点配置に近い配置であることを確認した。

第7章では、本論文の結論を述べ、今後の課題や展望について述べる。



# Sound Reproduction to a Limited Area by Multi-Point Control Method

Kazuya Yasueda

## Abstract

Virtual reality (VR) and augmented reality (AR) contents have been created because computers can process both a high-quality audio and visual. For such contents, since we must control the sound direction and its volume, headphones or earphones are generally for VR. If the sound can be controlled without relying on such devices, listener can be provided with better presence. Research has been conducted on sound field control using multiple loudspeakers to reproduce a recorded sound field in other locations or situations. Sound reproduction can be achieved for limited areas by applying these techniques.

Since people get specific information by sound reproduction to a limited area, these techniques are applied in automatic guidance systems at train stations and voice guidance systems at museums. Multilingual information and digital signage might eventually be reproduced by the different sounds to different areas. In this thesis, we study sound reproduction to a limited area using the multi-point control method with multiple loudspeakers.

Our thesis consists of the following seven chapters.

In chapter 1, we introduce and describe our study's purpose. We describe sound directivity control by a parametric loudspeaker and multiple dynamic loudspeakers as well as a sound field control method.

In chapter 2, we describe sound field control by multiple loudspeakers and sound

reproduction to a limited area by techniques that including a boundary surface control principle (BoSC) and a multi-point control method (MPCM), which controls the sound pressure for sound field control like the conventional method. MPCM matches the sound pressure inside a region between the target sound field and the reproduction area. In this study, we generate reproduction and suppression areas in a particular space using MPCM. In this chapter, we conducted an objective test by relative sound pressure level, a subjective test and a word intelligibility test. Our results showed that the content of the speech was understood outside the target area.

In chapter 3, we propose a sound reproduction method that combines MPCM and a signal decomposition method. We describe random vector decomposition, which is a conventional method for creating sub-signals, and propose a signal decomposition method using Linear Prediction Coding (LPC) residual signals. We experimented with a relative sound pressure level and a subjective test to evaluate our proposed method's performance.

In chapter 4, we propose a sound-masking method that addresses sound leakage of outside the target area. It is difficult to hear the target speech outside the target area because it is masked by the noise. However, since a slight masking noise is reproduced in the target area, we decreased its leakage in the target area by controlling it using MPCM and inverse signals. We subjectively evaluated the performance of our proposed method.

In chapter 5, we propose a control-point arrangement for sound reproduction to multiple areas because arranging the control points is important to reproduce sound to multiple areas with just a few loudspeakers. We studied the arrangement of the suppression control points inside the suppression area. We created a loudspeaker array for an experiment in real environments and compared the performance of the measured and free-field transfer function. We also conducted a word intelligibility test to evaluate the proposed arrangement's performance.

In chapter 6, we propose a method that determines the optimized control-point arrangement using the Genetic Algorithm. We improved the speech quality by introducing Perceptual Evaluation of Speech Quality (PESQ), which is an objective evaluation for the cost functions. We experimented by a computer simulation and identified the relationship between the relative sound pressure level and PESQ. The relative sound pressure level or PESQ improved. The obtained control-point arrangement approached our proposal in chapter 5.

In chapter 7, we describe our conclusions and future works.



# 目次

<b>第1章</b>	<b>序論</b>	<b>1</b>
1.1	研究背景 . . . . .	1
1.2	音の指向性制御 . . . . .	2
1.3	研究の目的 . . . . .	3
1.4	本論文の構成 . . . . .	3
<b>第2章</b>	<b>音場制御と指向性制御</b>	<b>5</b>
2.1	はじめに . . . . .	5
2.2	音場制御の原理 . . . . .	5
2.3	境界音場制御 . . . . .	7
2.4	多点制御法 . . . . .	8
2.4.1	多点制御法を用いた指向性制御によるエリア再生 . . . . .	9
2.4.2	重み付き多点制御法 . . . . .	10
2.4.3	拘束条件付き多点制御法 . . . . .	10
2.5	直線スピーカアレーの種類 . . . . .	11
2.6	エリア再生の評価方法 . . . . .	12
2.6.1	エリア再生のシミュレーション . . . . .	12
2.6.2	相対音圧レベル . . . . .	13
2.6.3	了解度 . . . . .	13
2.6.4	MOS 評価法 . . . . .	13
2.6.5	DMOS 評価法 . . . . .	14
2.7	評価実験 . . . . .	15

2.7.1	スピーカアレーと制御点の配置	15
2.7.2	実験結果	17
2.8	まとめ	19
<b>第3章</b>	<b>信号分解法と組み合わせたエリア再生</b>	<b>21</b>
3.1	はじめに	21
3.2	多点制御法によるサブ信号の再生	21
3.3	信号分解法によるサブ信号の生成	22
3.3.1	ランダムベクトル分解法	23
3.3.2	ノイズ駆動分解法	24
3.3.3	残差ランダムベクトル分解	26
3.3.4	雑音付加によるサブ信号の安定化	26
3.4	評価実験	28
3.4.1	客観評価実験	28
3.4.2	主観評価実験	30
3.5	まとめ	32
<b>第4章</b>	<b>目的方向外の音声に対する環境音によるマスキング手法</b>	<b>33</b>
4.1	はじめに	33
4.2	スピーチプライバシー保護のためのマスキング手法	33
4.2.1	パワーマスキング	34
4.2.2	情報マスキング	35
4.3	多点制御法による環境音の指向性制御	35
4.3.1	逆位相音の再生	37
4.3.2	環境音再生フィルタの分割	38
4.4	客観評価実験	38
4.5	主観評価実験	41
4.5.1	実験条件	41

4.5.2	実験結果	41
4.6	まとめ	45
<b>第5章</b>	<b>複数エリア再生</b>	<b>47</b>
5.1	はじめに	47
5.2	複数エリア再生	47
5.3	制御点配置の検討	48
5.4	計算機シミュレーション	51
5.4.1	実験条件	51
5.4.2	実験結果	51
5.5	実環境実験	56
5.5.1	実験条件	56
5.5.2	伝達関数の違いによる音圧の変化	56
5.5.3	単語了解度試験	59
5.6	まとめ	62
<b>第6章</b>	<b>遺伝的アルゴリズムを用いた制御点配置の決定法</b>	<b>63</b>
6.1	はじめに	63
6.2	遺伝的アルゴリズム	63
6.3	制御点の遺伝子表現	64
6.4	遺伝的アルゴリズムによる制御点配置の最適化	64
6.4.1	個体の選択方法	64
6.4.2	交叉と突然変異	65
6.4.3	評価関数	67
6.5	計算機シミュレーション	68
6.5.1	実験条件	68
6.5.2	突然変異確率の検討	69
6.5.3	実験結果	71

6.6 おわりに . . . . .	76
<b>第7章 結論</b>	<b>77</b>
7.1 本論文のまとめ . . . . .	77
7.2 今後の課題 . . . . .	79
<b>謝辞</b>	<b>81</b>
<b>参考文献</b>	<b>83</b>
<b>研究業績</b>	<b>91</b>

# 目 次

1.1	本論文の構成	4
2.1	任意の音場	6
2.2	境界音場制御の概要	7
2.3	直線スピーカアレーの種類	11
2.4	エンドファイア型アレーを用いたエリア再生における制御点配置	16
2.5	制御点上での相対音圧レベルの結果	17
2.6	各方向の単語理解度試験の結果	18
2.7	各方向での DMOS 評価法の結果	18
3.1	信号分解法と多点制御法によるエリア再生	22
3.2	ノイズ駆動分解法の処理の流れ	26
3.3	残差信号へのランダムベクトル分解の適用	27
3.4	ガウス雑音の付加によるサブ信号の安定化	28
3.5	各信号分解手法の相対音圧レベル	29
3.6	各信号分解手法の DMOS 評価結果	31
3.7	ガウス雑音の比率係数の違いによる DMOS 値の変化	31
4.1	マスキング手法を用いた多点制御法の制御点配置	36
4.2	一括方式で再生した環境音の相対音圧レベル	36
4.3	各方向の周波数特性 (黒：拘束条件変更前 赤：拘束条件変更後)	37
4.4	環境音を分割再生するときの制御点配置	38
4.5	環境音の相対音圧レベル	39
4.6	180 deg において目的音声に対して-10 dB としたときの各方向の SNR	40

4.7	DMOS 評価による主観評価実験結果 ( $\text{SNR}_{180} = -10$ )	42
4.8	DMOS 評価による主観評価実験結果 ( $\text{SNR}_{180} = -5$ )	43
4.9	DMOS 評価による主観評価実験結果 ( $\text{SNR}_{180} = 0$ )	44
5.1	複数エリア再生のイメージ	48
5.2	制御するエリアと再生制御点の配置	49
5.3	Area A に再生を行うときの制御点の配置	50
5.4	各配置の相対音圧レベル (抑圧制御点間隔 0.1 m)	52
5.5	各配置の相対音圧レベル (抑圧制御点間隔 0.02 m)	53
5.6	各配置の PESQ (抑圧制御点間隔 0.1 m)	54
5.7	各配置の PESQ (抑圧制御点間隔 0.02 m)	55
5.8	実験を行った防音室の見取り図	57
5.9	作成した 16ch 直線スピーカアレー	57
5.10	Area A に再生を行った時の相対音圧レベル	58
5.11	Area B に再生を行った時の相対音圧レベル	58
5.12	単語了解度試験時の受聴者の位置	60
5.13	受聴位置 A での単語了解度試験結果	61
5.14	受聴位置 B での単語了解度試験結果	61
6.1	遺伝的アルゴリズムの処理の流れ	65
6.2	$n = 3$ としたときの一様交叉の例	66
6.3	各エリアとスピーカアレーの配置	66
6.4	音圧と PESQ の測定点	69
6.5	突然変異確率を変化させたときの収束世代数と適合度	70
6.6	PESQ の重み係数による相対音圧レベルと PESQ の平均	71
6.7	各世代ごとの相対音圧レベルと PESQ の変化 ( $A = 0$ )	72
6.8	各世代ごとの相対音圧レベルと PESQ の変化 ( $A = 5$ )	72
6.9	各世代ごとの相対音圧レベルと PESQ の変化 ( $A = 10$ )	72

6.10 各世代の制御点配置と最終世代の相対音圧マップ ( $A = 0$ ) . . . . .	73
6.11 各世代の制御点配置と最終世代の相対音圧マップ ( $A = 5$ ) . . . . .	74
6.12 各世代の制御点配置と最終世代の相対音圧マップ ( $A = 10$ ) . . . . .	75



# 表 目 次

2.1	MOS 評価実験の評価カテゴリ . . . . .	14
2.2	DMOS 評価実験の評価カテゴリ . . . . .	15
2.3	聞き取り難さと不快さの評価カテゴリ . . . . .	15
5.1	各配置の制御点数 (再生制御点を含む) . . . . .	50



# 第1章 序論

## 1.1 研究背景

近年、計算機の性能向上により、高解像度な映像や臨場感のある音響を処理できるようになり、VR(Virtual Reality)やAR(Augmented Reality)などで多くのコンテンツが作成されている [1, 2, 3]. その際、音の方向や大きさを制御することは臨場感を与える重要な要素となる. 現状, VRなどではヘッドホンやイヤホンなどを用いて個別に音を再生することが主流であるが, これらを使用せずに音を制御することができれば, より高臨場感を与えられることが期待できる [4, 5, 6]. そこで, 複数のスピーカを用いて音場を制御する音場制御の研究が行われている [7, 8, 9, 10, 11]. これらの研究は, 収録した音場を別の領域で再現することを目的としているが, この理論を応用することで音を特定の方向や領域のみに再生する, 音のエリア再生を行うことができる.

音のエリア再生は, 特定の情報を必要な人のみにピンポイントで届けることができるため, 駅ホームでの案内放送や美術館・博物館などでのガイダンスシステムなどで利用することができる. また, 複数の場所に異なる音声を再生する複数エリア再生 [12, 13, 14, 15, 16] は多言語による案内放送やデジタルサイネージへの応用が期待できる.

特定の人へのピンポイントな情報伝達は, その情報を必要とする人以外に対しては騒音となり得ることから, 高い指向性が必要である. また, スピーチプライバシーの観点から音声内容が周囲へ漏えいしないことが重要である [17, 18]. 一方, 複数エリア再生では複数人での利用が想定されるため, ある程度広い範囲に再生されることが望ましい [19].

## 1.2 音の指向性制御

音の指向性を制御する方法は大きく分けて2つある。1つはスピーカ自体の再生方式を工夫し、単体のスピーカで指向性を持たせる方法である。もう1つは、通常のダイナミックスピーカを複数個用いて音場を制御する方法である。

単体のスピーカシステムで指向性を制御する技術として、超音波と空気の非線形性を用いたパラメトリックスピーカがある [20, 21, 22, 23]。パラメトリックスピーカは複数の超音波エミッタを用いて、周波数の接近した2つの超音波(一次波)を大きな振幅で放射することで、その差分を周波数とする振幅変調波(二次波)が発生する。一次波は球面波として減衰するが、二次波は一次波の非線形相互作用により、減衰が遅く、鋭い指向性を持つ。したがって、二次波を可聴音として振幅変調すれば、入力信号が空間上で復調しエリア再生を実現できる。しかし、パラメトリックスピーカは高調波歪みの発生や復調距離によっては音声品質が十分でない問題がある [23]。さらに、大音圧で超音波を放射することから人体に影響を及ぼす超音波ばく露が懸念されている [24]。

一方、複数のスピーカを用いる技術は、各スピーカから再生される音の振幅と位相を制御することで任意の指向特性を実現する。その中で、最も単純な方法が遅延和法 [25] である。遅延和法は再生したい方向に複数のスピーカの位相を合わせることで、指向性を持たせる手法である。一般的にホールやライブ会場などで用いられるラインスピーカアレーと呼ばれるシステムは、直線上に複数のスピーカを並べることで、信号処理を必要とせずにスピーカアレーに対して垂直方向への指向性を実現する。また、各スピーカの遅延量を制御するフィルタを用いることで、任意の方向への指向特性を実現できる。

遅延和法は位相を揃えることで特定方向での強調を行うことができるが、それ以外の音を抑圧したい方向に対して、陽に制御を行わないため、エリア再生には向いていない。また、低域での指向特性を形成しにくい [26]。そこで、音場制御の理論に基づいた制御手法が提案されている。空間フーリエ変換を用いて波数領域で制御を行う方法 [14, 16, 27] や円調和関数展開 [28, 29]、球面調和関数展開を用いる手法 [30]、時間領域で音圧を直接制御する手法 [31, 32, 33, 34, 35] などがある。空間フーリエ変換を用いる

手法や球面調和関数展開を用いる手法はスピーカアレーや制御領域の形状に関して制約を設けることで、解析的にスピーカへの入力信号を求めることができる。一方で、制御領域の音圧を直接制御する手法は、スピーカアレーや制御領域の形状に制約はないが、スピーカと制御点間の伝達関数に基づき逆フィルタによって制御を行うため、フィルタが不安定になる場合がある [36, 37].

### 1.3 研究の目的

本研究の目的は、多点制御法を用いて任意の場所へのピンポイントな再生とある程度広い範囲への再生を実現することである。

一般的に複数のスピーカを用いて音場を制御する手法は、スピーカ数が多いほど制御性能が良い。しかし、スピーカ数が十分でない場合エリア再生の性能が低下するため、結果として音声内容が漏えいする問題がある。そこで本研究では、多点制御法によるエリア再生手法において、エリア外の漏えいを考慮したエリア再生手法を提案する。また、1つのスピーカアレーで複数の音声を異なるエリアに再生を行うための制御点配置について提案する。

### 1.4 本論文の構成

図 1.1 に本研究の概要を示す。第 2 章では、複数のスピーカアレーによる音場制御手法について説明し、多点制御法による指向性制御手法の概要について述べる。第 3 章と第 4 章では、狭い範囲へのエリア再生を実現するために、目的のエリア外に漏えいする音声を聞き取れなくする手法を検討する。第 3 章では、信号分解法を用いて入力信号を複数のサブ信号に分解し、空間上で再合成することでエリア再生を実現する手法を提案する。第 4 章では、多点制御法による指向性制御が不十分な場合において、環境音によるマスキング手法を用いることで聞き取れる範囲を制御する手法を提案する。第 5 章と第 6 章では広い範囲での複数エリア再生を実現するために、抑圧を行う制御点配置について検討を行う。第 5 章では、複数の領域に異なる音声を再生するための

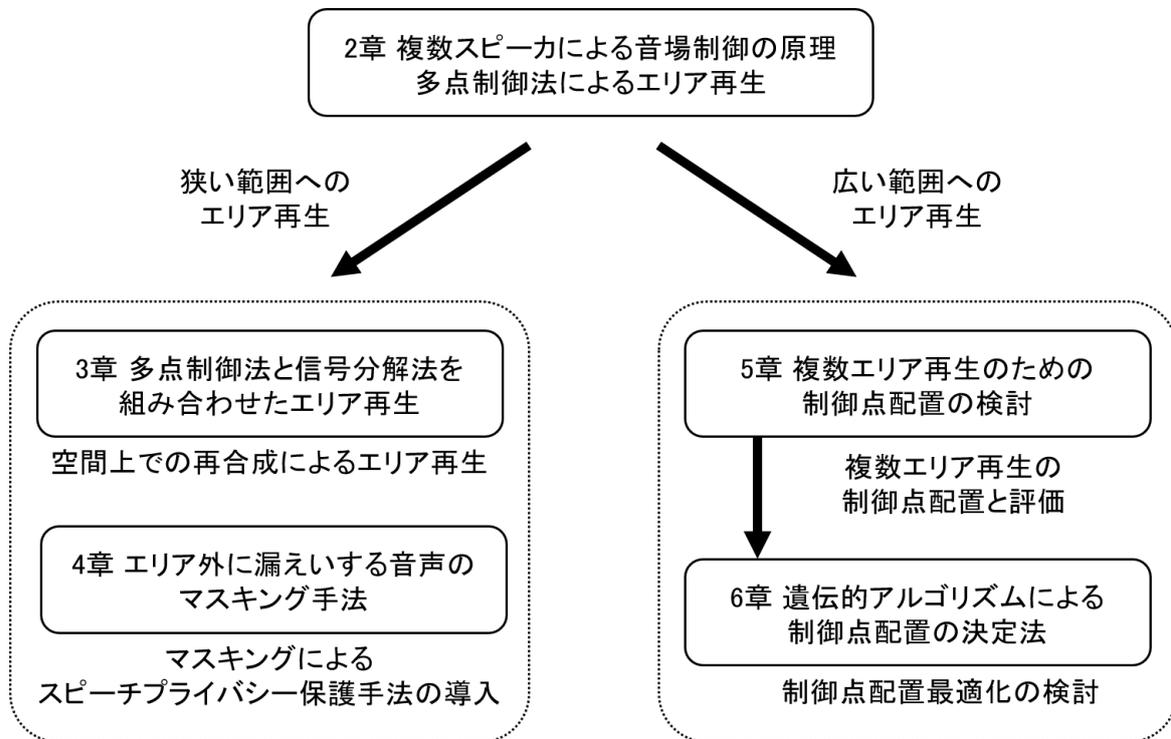


図 1.1: 本論文の構成

制御点配置についてL字型の抑圧制御点配置を提案する。また、実環境での評価実験の結果を述べる。第6章では、遺伝的アルゴリズムを用いて最適な制御点配置を決定する手法を提案する。最後に第7章では、本論文の結論を述べる。

## 第2章 音場制御と指向性制御

### 2.1 はじめに

本章では、複数のスピーカを用いて音場を制御する手法とその原理を応用したエリア再生の手法について述べる。音場を制御する目的として主に音場再現と指向性制御がある。音場再現は任意の音場を収録し、別の空間で再現する技術であり、パーソナルオーディオやシアターなどで臨場感のある音場を生成する際に用いられる。一方、指向性制御手法は特定の方向や場所といった限られた空間のみに音を再生する技術である。

以下、2.2節において、音場制御の基礎となる原理について述べる。2.3節において、境界音場制御と逆フィルタによる音場制御法について述べる。2.4節において、多点制御法の原理を説明し、多点制御法を用いた指向性制御によるエリア再生手法について述べる。2.5節において、直線スピーカアレーの種類と特性について述べる。2.6節で多点制御法を用いたエリア再生手法の性能評価実験について述べる。2.8節において、本章のまとめを述べる。

### 2.2 音場制御の原理

図 2.1 に示す 3次元領域  $V$  の内部音場  $P(r)$  を領域  $V$  の境界  $S$  上に配置したスピーカを用いて所望の音場と一致するように制御する。境界  $S$  上の位置ベクトルを  $r_s$  とし、 $r_s$  におけるスピーカの駆動信号を  $D(r_s)$  とすると、領域内部の音圧  $P(r)$  は以下のようなになる [36]。

$$P(r) = \int \int_S D(r_s) G(r - r_s) dS \quad (2.1)$$

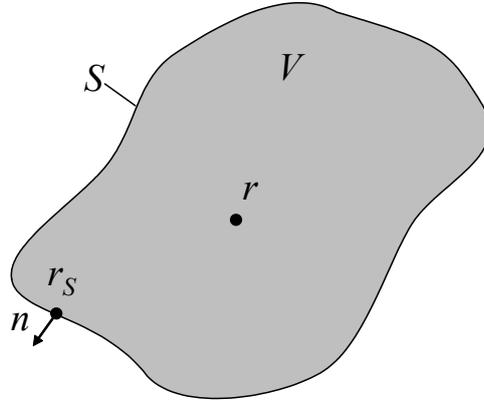


図 2.1: 任意の音場

ここで、 $G(r - r_s)$  は位置  $r_s$  のスピーカから位置  $r$  までの伝達特性を表す。この式は領域内部の音圧はスピーカから再生される信号と、スピーカから任意の点までの伝達関数の畳み込みとなっていることがわかる。

一方、音の伝搬や回折を説明したホイヘンス-フレネルの原理がある。この原理は、波面は前時刻の波面上に想定した二次波源からそれぞれ波面が生成され、その二次波面の包絡によって形成されると考えている。この原理について、キルヒホッフはヘルムホルツ方程式の近似形式として見なせることを示した [38]。この式はキルヒホッフ-ヘルムホルツ積分方程式と呼ばれ、ホイヘンス-フレネルの原理を数学的に記述した式であり、波動方程式の周波数表現にグリーンの定理を適用することで導出できる。式 (2.2) にキルヒホッフ-ヘルムホルツ積分方程式を示す。

$$P(r) = \int \int_S \left( \frac{\partial P(r_s)}{\partial n} G_{3D}(r - r_s) - P(r_s) \frac{\partial G_{3D}(r - r_s)}{\partial n} \right) dS \quad (2.2)$$

ここで、 $G_{3D}(r - r_s)$  は 3次元自由空間グリーン関数であり、 $n$  は境界  $S$  の法線方向ベクトルを表す。 $G_{3D}(r - r_s)$  はモノポール特性の音源、 $\partial G_{3D}(r - r_s)/\partial n$  はダイポール特性の音源からの伝達関数である。この式から、境界  $S$  上に設置したモノポール特性とダイポール特性の音源を用いて、所望の音場となるように音圧と音圧勾配を制御す

ることで、音場全体の制御が可能であることがわかる。また、境界  $S$  上の音圧  $P(r_s)$  と音圧勾配  $\partial P(r_s)/\partial n$  によって、音場が一意に決定できる。ここで、音圧勾配とは境界  $S$  における法線方向の音圧の変化量である。

キルヒホッフ-ヘルムホルツ積分方程式によって、任意の領域内の音を制御するためには、その境界上において音圧と音圧勾配を制御すればよいことがわかった。しかし、一般的なマイクロホンやスピーカでは、音圧勾配の収録やダイポール特性を実現することは困難である。特に音圧勾配の計測については、2つのマイクロホンを近接して設置し、その差分から音圧勾配を近似する2マイクロホン法 [39] などがあるが、3次元領域のすべての境界で測定することは困難である。

## 2.3 境界音場制御

キルヒホッフ-ヘルムホルツ積分方程式は、領域の境界において音圧と音圧勾配を制御することで音場を再現できることを示している。したがって、必ずしも境界上に音源を置く必要はなく、図 2.2(a) のように領域外部の任意の位置に設置し、境界上で音圧と音圧勾配が等しくなるように制御すればよい。このとき、式 (2.2) の  $G_{3D}(r - r_s)$  と  $\partial G_{3D}(r - r_s)/\partial n$  は境界と音源の位置関係によって決まる定数とみなすことで音源位置に対する制約がなくなる。この原理を用いて音場を制御する手法が境界音場制御

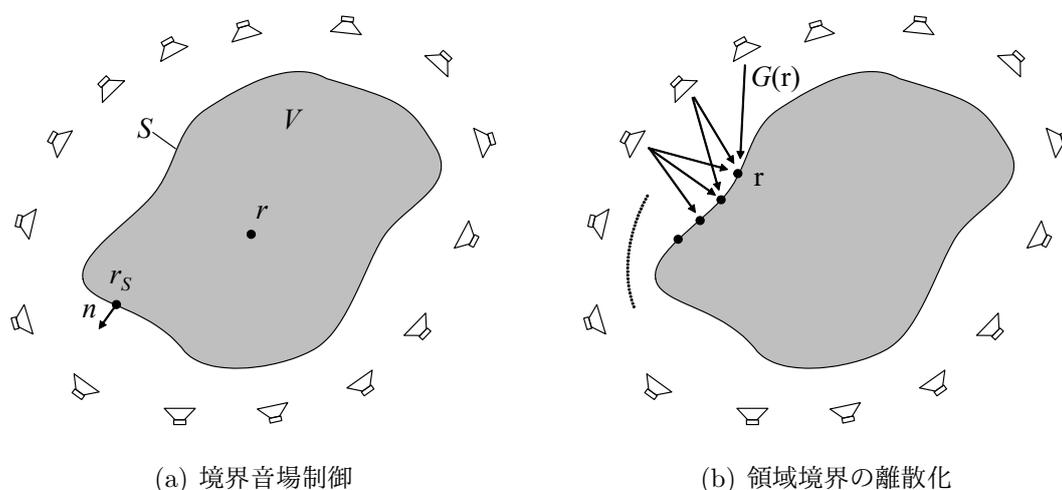


図 2.2: 境界音場制御の概要

である [40, 41, 42].

境界音場制御では積分方程式をディリクレ問題と考えることで、固有周波数以外は境界上で音圧か音圧勾配のどちらかを制御することによって領域内部の制御が可能であることを示している。これは、収録・再生が困難である音圧勾配を用いることなく、境界上の音圧のみを制御することで、音圧勾配も同時に制御できることを意味している。また、境界音場制御では領域境界の音圧と音圧勾配について、境界要素法における積分方程式の離散化と同様の方法を用いることで境界を離散化している。

再生する音源が既知の場合、図 2.2(b) のように領域外部に設置したスピーカと領域境界の離散化された点 (制御点) までの伝達関数  $G(r)$  を計測し、制御点での音圧が測定した原音場の音圧と一致するような逆フィルタを設計すればよい。したがって、スピーカと制御点間の伝達関数を 1 度測定し、逆フィルタを求めておけばどのような音源に対しても適用できる [43].

## 2.4 多点制御法

境界音場制御では、領域境界において音圧と音圧勾配を制御することを考えていたが、領域内部の音圧を直接一致させることで音場再現を行う多点制御法 (Pressure-Matching 法) がある [32, 33]. 式 (2.3) のように合成した音場と所望の音場の誤差が最小となるように  $m$  番目のスピーカに対応するフィルタ係数  $w_m$  を求めることで領域内の音圧を制御する。

$$\min_{w_m} \int_{r \in V} \left( \sum_{m=1}^M w_m g_m(r) - d(r) \right)^2 dr \quad (2.3)$$

$g_m(r)$  はスピーカ  $m$  から位置  $r$  までの伝達関数、 $d(r)$  は所望の音場である。しかし、この式は領域  $V$  に関する積分を含むため、直接最適化することは困難である。そこで、領域を離散化し、離散的な制御点での音圧を一致させる。離散化することで、式 (2.3) は式 (2.4) のようになる。

$$\mathbf{W} = \arg \min_{\mathbf{w}} \|\mathbf{GW} - \mathbf{D}\|^2 \quad (2.4)$$

$$\mathbf{G} = \begin{bmatrix} G_{11}(\omega, r) & G_{12}(\omega, r) & \cdots & G_{M1}(\omega, r) \\ G_{12}(\omega, r) & G_{22}(\omega, r) & \cdots & G_{M2}(\omega, r) \\ \vdots & \vdots & \ddots & \vdots \\ G_{1N}(\omega, r) & G_{2N}(\omega, r) & \cdots & G_{MN}(\omega, r) \end{bmatrix} \quad (2.5)$$

$$\mathbf{W} = [w_1(\omega, r), w_2(\omega, r), \dots, w_M(\omega, r)]^T \quad (2.6)$$

$$\mathbf{D} = [d_1(\omega, r), d_2(\omega, r), \dots, d_N(\omega, r)]^T \quad (2.7)$$

ここで、 $\omega$  は周波数、 $\mathbf{W}$  はフィルタ係数ベクトル、 $\mathbf{G}$  はスピーカと制御点間の伝達関数行列、 $\mathbf{D}$  は制御点での所望の音圧特性を要素に持つベクトルである。式 (2.4) は式 (2.8) のように一般化逆行列による最小二乗法としてフィルタ係数を求めることができる [44]。

$$\mathbf{W} = (\mathbf{G}^H \mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{G}^H \mathbf{D} \quad (2.8)$$

$\lambda$  は正則化パラメータ、 $\mathbf{I}$  は単位行列、 $(\cdot)^H$  は複素共役転置を表す。多点制御法による音場再現は、領域内において制御点を密に設置する必要があるが、伝達関数  $\mathbf{G}$  の独立性が低くなるとフィルタが不安定になる場合がある [36, 37]。また、正則化パラメータは大きく設定すると外部からの雑音に頑健になり安定した計算ができるが、音場制御の精度が低下する。

### 2.4.1 多点制御法を用いた指向性制御によるエリア再生

多点制御法を用いて音場再現を行う場合、式 (2.8) の  $\mathbf{D}$  に収録した音場の音圧ベクトルを代入し、フィルタ係数を求めることで音場の再現が可能である。したがって、 $\mathbf{D}$  の要素  $d_n(\omega, r)$  に対して 0 を代入することで、その制御点での音圧をゼロにすることができ、1 を代入することで入力信号をそのまま再生することができる。これを利用し、式 (2.9) のように音を再生したい制御点の音圧特性を 1、抑圧したい領域の制御点の音圧特性を 0 とすることで、音圧特性を 1 とした制御点へ指向性が生じ、エリア再生が

実現できる [45, 46].

$$d_n(\omega) = \begin{cases} 1 & \text{Reproduction point} \\ 0 & \text{Suppression point} \end{cases} \quad (2.9)$$

本論文では、再生を行う  $d_n = 1$  の制御点を再生制御点 (Reproduction control point), 抑圧を行う  $d_n = 0$  の制御点を抑圧制御点 (Suppression control point) と呼ぶ.

## 2.4.2 重み付き多点制御法

多点制御法ではスピーカ数より制御点数が多い場合, 式 (2.8) のようにフィルタ係数を最小二乗法によって求める. しかし, エリア再生を行う場合はスピーカの位置や部屋の形状などにより, 抑圧量を大きくしたい場所とそうでない場所が存在する可能性がある. そこで, 重み付き最小二乗法を用いて, 制御点に重みをつけることで制御領域内での抑圧量の調整を可能とする. 式 (2.10) に重み付き最小二乗法を用いた多点制御法を示す.

$$\mathbf{W} = (\mathbf{G}^H \mathbf{A} \mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{G}^H \mathbf{A} \mathbf{D} \quad (2.10)$$

$$\mathbf{A} = \text{diag}[a_1, a_2, \dots, a_N] \quad (2.11)$$

ここで,  $a_n$  は各制御点での重み係数である.  $a_n$  がすべて 1 の時は式 (2.8) と一致する. 制御点ごとに重みを設定することで抑圧量を制御することができる.

## 2.4.3 拘束条件付き多点制御法

多点制御法では周波数ごとに独立したフィルタ係数を計算する. 式 (2.5) の伝達関数行列は周波数によって逆行列計算の精度が異なるため, 結果として再生制御点での音質が低下する可能性がある. そこで, 再生制御点での音質を補正するフィルタ  $\mathbf{F} = [f_1, f_2, \dots, f_m]$  を求める.

$$\mathbf{F} = (\mathbf{G}_{\text{rep}} \mathbf{W})^{-1} \quad (2.12)$$

ここで、 $\mathbf{G}_{\text{rep}}$  は再生制御点に対応する伝達関数行列である。補正フィルタをそれぞれ式 (2.8) のフィルタ係数にかけることで、再生制御点での音質を保証することができる。この補正フィルタは逆フィルタとなっているため、再生制御点をスピーカアレーから遠くに設置するなど、伝達関数が非常に小さい値をとる場合はフィルタが発散する可能性がある。

## 2.5 直線スピーカアレーの種類

多点制御法はスピーカと制御点間の伝達関数によって制御を行うため、スピーカや制御点を自由に配置することができる。音場全体を制御する場合、スピーカは制御領域の周囲に配置することが望ましい。一方でエリア再生の場合は、デジタルサイネージでの利用や駅ホームでの利用を考えると、前方への放射は許容される場合が多い。そこで、本研究ではスピーカを一定間隔で直線状に配置した直線スピーカアレーを用いる。

図 2.3 に 2 種類の直線スピーカアレーを示す。図 2.3(a) はスピーカアレーの軸と同じ方向に再生する方式で、エンドファイヤ型アレーと呼ばれる。一方、図 2.3(b) はスピーカアレーの軸に垂直な方向へ再生する方式で、ブロードサイド型アレーと呼ばれる。スピーカを点音源と仮定すると、スピーカと制御点間の伝達関数は距離の関数となるため、スピーカアレーを円筒と見たときの半径方向は同一の指向特性となる。し

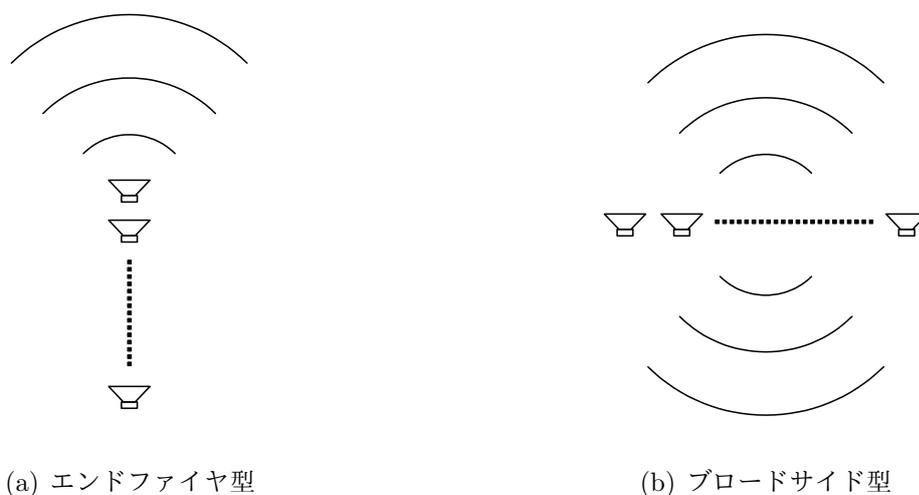


図 2.3: 直線スピーカアレーの種類

たがって、スピーカが存在する平面状のみを考えた場合、ブロードサイド型アレーはスピーカの後方においても前方と同様の指向特性となる。

エンドファイヤ型アレーは空間上の1点に再生する場合に制御効果が高く、ブロードサイド型アレーは広い領域を制御する場合に用いられる。前述のようにブロードサイド型アレーは後方にも再生音が放射されるため、壁に埋め込むなど後方からの反射対策を行う必要がある。

## 2.6 エリア再生の評価方法

エリア再生の性能は、再生対象となる領域が他の領域と比べてどの程度音圧が大きいかによって評価を行う。また、領域内に再生した音声について聞き取りの容易さや品質を評価する必要がある。本節では、エリア再生の性能や再生品質を評価するための尺度や評価手法について述べる。

### 2.6.1 エリア再生のシミュレーション

本研究では反射や残響のない自由空間を仮定し、シミュレーションを行う。多点制御法はスピーカと制御点間の伝達関数に基づいてフィルタを設計するため、式(2.5)の伝達関数に式(2.13)で表される自由空間伝達関数を用いる。

$$G_{mn}(\omega) = \frac{1}{4\pi r_{mn}} \exp(-j\omega \frac{r_{mn}}{c}) \quad (2.13)$$

ここで、 $r_{mn}$  はスピーカ  $m$  から制御点  $n$  までの距離、 $c$  は音速である。また、任意の場所での再生音声に関しても多点制御フィルタを畳み込んだ入力信号に対して、式(2.13)の伝達関数を畳み込むことで生成することができる。

## 2.6.2 相対音圧レベル

相対音圧レベル (Relative Sound Pressure Level) はある点を基準とした音圧比であり, エリア再生の性能を客観的に評価する. 任意の位置  $r$  での相対音圧レベルは式 (2.14) で表される. 値が大きいほど, その場所での音圧が大きく, 小さいほど音圧が小さいことを示す.

$$\text{Relative SPL}(r) = 10 \log_{10} \frac{P(r)}{P(r_{ref})} \quad (2.14)$$

ここで,  $P(r)$  は点  $r$  でのパワー,  $P(r_{ref})$  は基準点のパワーである.

## 2.6.3 了解度

了解度は意味のある単語や文章を聴取し, その正解率によって音声品質を評価を行う主観評価の一つである. 単語の聞き取りを単語了解度, 文章の聞き取りを文章了解度という. 了解度試験は試験が難しく, 長時間に及ぶことから一般的に単語了解度試験が用いられることが多い. 単語了解度試験に用いる音声は, 天野らが開発した単語了解度試験用音声データセット (FW03, FW07) が用いられる [47]. この音声セットは単語の親密度別になっており, 本研究では FW03 のうち親密度が最も高い, 親密度 5.5~7.0 の音声セットを用いる. 音声は話者 4 名 (男女各 2 名) が発話した 4 mora の単語である. また, 回答は完全一致で正解とし, その正解率を単語了解度とする.

## 2.6.4 MOS 評価法

MOS 評価法は対象音声を 1 度だけ聴取し, その音質について 5 段階で評価する主観評価手法である [48]. 表 2.1 に MOS 評価に用いる一般的な評価カテゴリを示す. 電話などの音声品質を評価する MOS 評価法は, ITU-T P. 800 で定義されており, 絶対範疇尺度法を用いて「非常に良い」から「非常に悪い」の 5 段階で絶対評価を行う.

表 2.1: MOS 評価実験の評価カテゴリ

評点	評価カテゴリ
5	非常に良い
4	良い
3	普通
2	悪い
1	非常に悪い

### 2.6.5 DMOS 評価法

DMOS 評価法は、先にリファレンス音声 (処理前の原信号) を聴取した後、評価対象となる音声を聴取し、MOS 評価法と同様に 5 段階で評価する主観評価手法である [48]. 表 2.2 に DMOS 評価に用いる一般的な評価カテゴリを示す. 被験者はリファレンス音と比較して対象音声を評価するため、MOS 評価法と比べて劣化に敏感な評価をすることができる. しかし、リファレンス音を聴取する必要があるため、MOS 評価法と比べて約 2 倍の試験時間を要する.

本研究では、音声の劣化に加えて、表 2.3 の評価カテゴリを用いて聞き取り難さ (Difficulty of Speech Listening) と不快さ (Harshness) について評価を行う. 値が大きいほどポジティブな評価、小さいほどネガティブな評価とし、5 段階で評価する.

表 2.2: DMOS 評価実験の評価カテゴリ

評点	評価カテゴリ
5	劣化が全く認められない
4	劣化が認められるが気にならない
3	劣化がわずかに気になる
2	劣化が気になる
1	劣化が非常に気になる

表 2.3: 聞き取り難さと不快さの評価カテゴリ

評点	聞き取り難さ	不快さ
5	全く聞き取り難くない	全く不快ではない
4	やや聞き取り難い	それほど不快ではない
3	聞き取り難い	多少不快である
2	とても聞き取り難い	だいぶ不快である
1	聞き取れない	非常に不快である

## 2.7 評価実験

エンドファイヤ型アレーを用いて空間上の1点に再生する場合のエリア再生の性能について、計算機シミュレーションによる相対音圧レベルの算出と単語了解度試験、DMOS 評価試験を行った。本節では実験条件と結果について述べ、多点制御法の問題点を示す。

### 2.7.1 スピーカアレーと制御点の配置

図 2.4 にスピーカアレーと制御点の配置を示す。図 2.4 のように、再生制御点は1点とし、スピーカアレーの軸と同一の方向に設置する。エンドファイヤ型スピーカアレーはアレーの軸に対称になっているため、抑圧制御点はスピーカアレーを囲うように半円状に設置する。

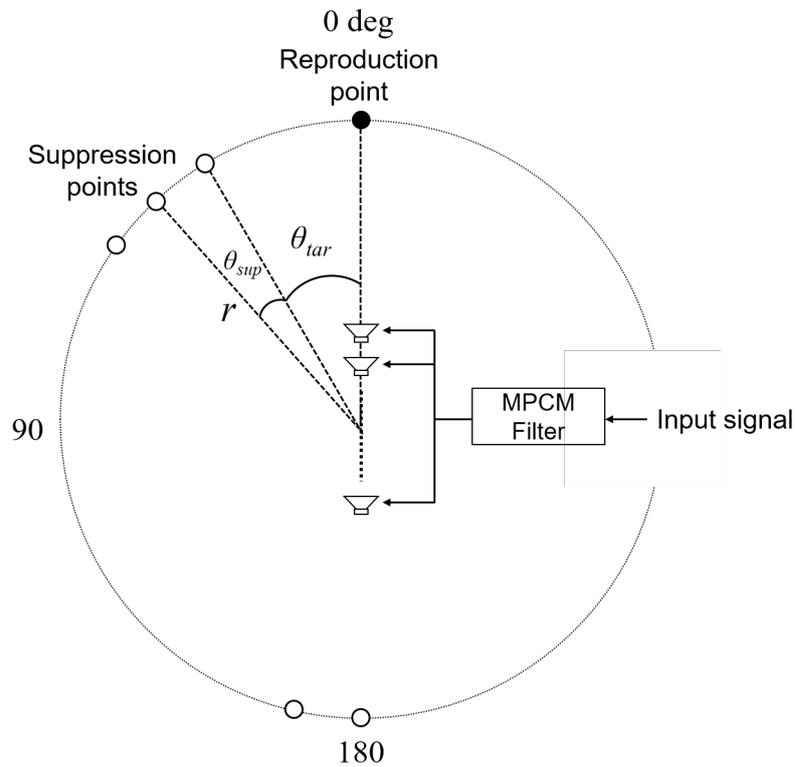


図 2.4: エンドファイア型アレーを用いたエリア再生における制御点配置

スピーカ数は 8 ch とし，制御点配置は図 2.4 における  $\theta_{tar} = 5 \text{ deg}$ ， $\theta_{sup} = 5 \text{ deg}$ ， $r = 1 \text{ m}$  とした．制御点数は 37 点である．使用した音源は 100~7000 Hz に帯域制限した男女の発話音声で，サンプリングレートは 16 kHz である．

相対音圧レベルは男女各 5 音声，計 10 音声の平均値である．単語理解度試験と DMOS 評価試験の被験者数は 20 代の男女 12 名，評価した音声数は各方向それぞれ 20 音声である．評価はシミュレーションによって作成した各方向の音声をヘッドホンを用いて聴取させた．このとき，0 deg 方向での音圧を約 60 dBA とし，90，180 deg 方向の音声は音圧の調整をせずに，多点制御法によって音圧が低下した状態で聴取させた．

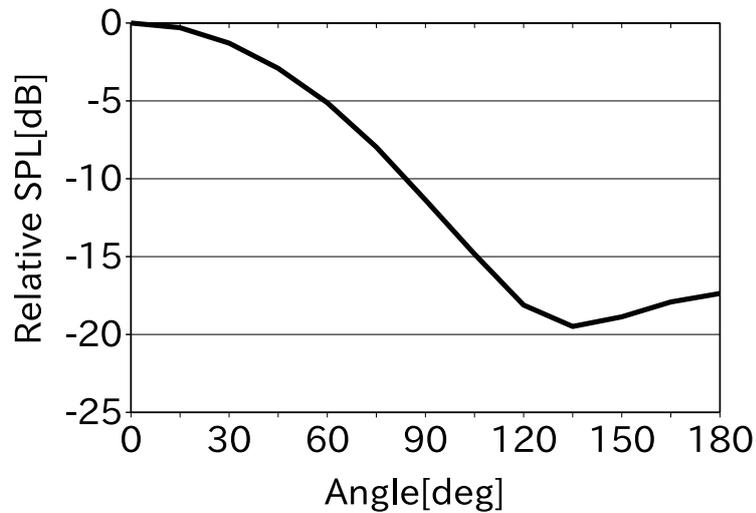


図 2.5: 制御点上での相対音圧レベルの結果

## 2.7.2 実験結果

図 2.5 に抑圧制御点上での相対音圧レベル，図 2.6 に単語了解度試験の結果，図 2.7 に DMOS 評価の結果を示す．図 2.7 のエラーバーは 95 %信頼区間を示す．図 2.5 から，目的方向外では目的方向から離れるにつれて再生音が抑圧されていることがわかる．特に，135 deg 方向では約 20 dB の抑圧が確認できる．しかし，図 2.6 からどの方向においても，単語了解度が 100 %となっている．また，図 2.7 から，90, 180 deg 方向での聞き取りやすさと不快さが 4 程度となっている．これらのことから，目的方向外において音圧を低下させることはできているが，完全にゼロになっておらず音声劣化も少ないため，音声内容を聞き取ることができると考えられる．

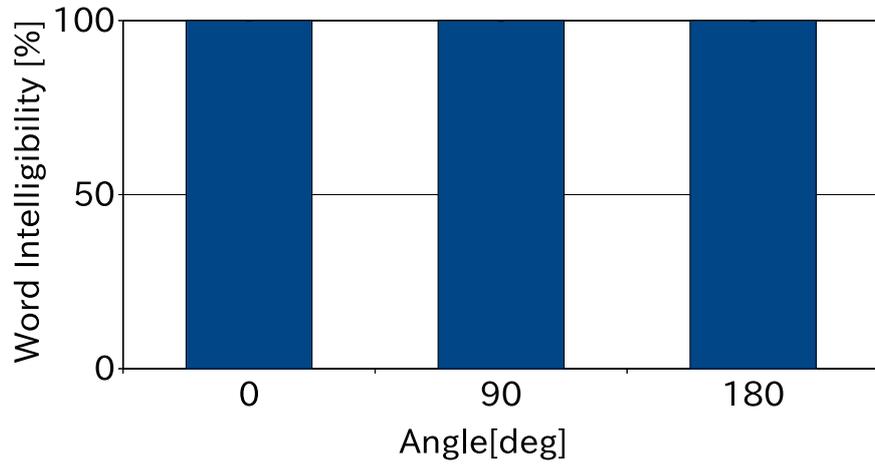


図 2.6: 各方向の単語了解度試験の結果

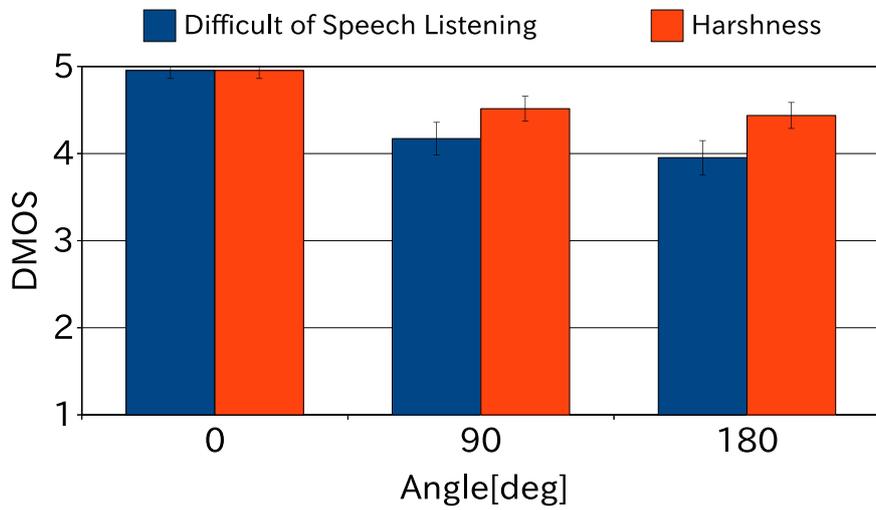


図 2.7: 各方向での DMOS 評価法の結果

## 2.8 まとめ

本章では，キルヒホッフ-ヘルムホルツ積分方程式に基づく複数スピーカによる音場制御の理論について述べ，領域内の音圧を直接制御する多点制御法について説明した．また，本研究で用いる直線スピーカアレーの種類について説明した．さらに，エリア再生の性能を評価する方法を説明し，多点制御法を用いたエリア再生手法について，計算機シミュレーションによる相対音圧レベルと単語了解度試験，DMOS評価試験を行った．実験結果から，目的方向外での音の抑圧は確認できたが，音声品質に劣化が少なく，音声内容が聞き取れる可能性を示した．



# 第3章 信号分解法と組み合わせた エリア再生

## 3.1 はじめに

多点制御法による指向性制御手法では対象領域の音圧を制御することで実現していた。しかし、スピーカ数や制御点数の物理的制約や最小二乗法の誤差から、音圧を完全にゼロにできず、目的方向外や対象エリア外に漏えいする問題があった。また、音声の品質劣化も少ないことから漏えいした音声の内容が聞き取れてしまうという問題があった。本章では、信号分解法を用いて、入力音声をサブ信号に分解し、多点制御法を用いて空間上で再合成することによって狭い範囲へのエリア再生を実現する手法を提案する。

以下、3.2節において、多点制御法を用いたサブ信号の再生方法について述べる。3.3節において、信号分解法の各手法とサブ信号の生成方法について述べる。3.4節において、信号分解法を用いたエリア再生の性能を評価するための評価実験について述べる。3.5節において、本章のまとめを述べる。

## 3.2 多点制御法によるサブ信号の再生

入力信号を何らかの処理によって複数の信号に分解する。このとき、分解した信号をすべて加算すると原信号に復元するように分解処理を行う。本研究ではこの分解した信号をサブ信号と呼ぶ。

生成したサブ信号を複数のスピーカを用いて再生することで、サブ信号同士の位相が一致する場所では原信号に復元し、それ以外の場所では位相がずれることで聞き取

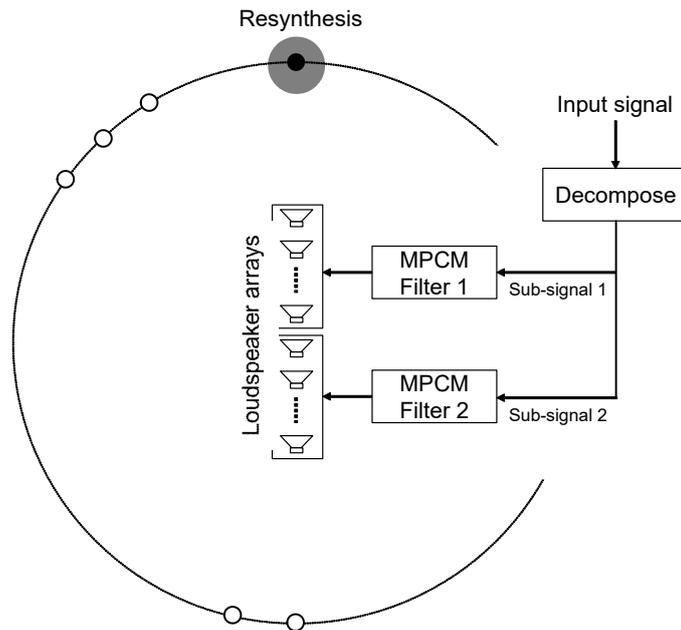


図 3.1: 信号分解法と多点制御法によるエリア再生

り難くなる。しかし，サブ信号自体が周囲に漏えいすることで不快さが上昇する可能性がある。そこで，多点制御法によって指向性の制御を行うことで周囲への漏えいを抑圧し，不快さを低減させる。図 3.1 に 2 つのサブ信号を用いたエリア再生の概要を示す。エンドファイア型の直線スピーカアレーを 2 つに分割し，それぞれで多点制御フィルタを設計する。このとき，再生を行う制御点を同じ位置に設定する。こうすることで，再生制御点ではサブ信号が復元することで原音声聞き取れ，抑圧する領域では多点制御法による音圧の低下とサブ信号の位相ずれによって聞き取り難くなることが期待できる。

### 3.3 信号分解法によるサブ信号の生成

信号分解法は，音声信号を何らかの線形処理によって複数のサブ信号に分解する手法である。線形処理によって分解するため，サブ信号をすべて加算すると元の音声に復元する。サブ信号単体では聞き取れないように分解を行い，そのサブ信号を個別に伝送することができれば，秘話性の高い情報伝達が可能となる。また，サブ信号をそ

それぞれ独立したスピーカから再生することで、空中で再合成し、特定の人に情報を伝達する手法も提案されている [49, 50]. サブ信号単体で聞いたときには元の音声内容が聞き取れないことが望ましいが、サブ信号が白色雑音に近いと再合成時のわずかな誤差によって元の信号に戻らず、再生音声の劣化や可聴範囲が狭くなることが考えられる. 本節ではサブ信号の生成方法として、ランダムベクトル分解法 [51] と LPC 残差信号を用いた分解法、それぞれを組み合わせた分解法について述べる.

### 3.3.1 ランダムベクトル分解法

入力信号をサブ信号に分解する手法としてランダムベクトル分解法がある [51]. ランダムベクトル分解法は音声信号をベクトルと見て成分分解を行い、任意の個数を合成し、サブ信号を生成する手法である. 音声信号は式 (3.1) のように任意の信号と係数の線形結合で表現することができる.

$$s(n) = \sum_{k=1}^K a_{rv}(k)v(k, n) \quad (3.1)$$

ここで、 $s(n)$  は入力信号、 $a_{rv}(k)$  は係数、 $v(k, n)$  はあらかじめ用意した任意の雑音信号、 $K$  は 1 フレームのデータ点数である. この式を満たすような  $a_{rv}(k)$  を求めることで音声信号を最大  $K$  個のサブ信号に分解することができる. 式 (3.1) を行列形式で表すと以下のようなになる.

$$\mathbf{s} = \mathbf{V}\mathbf{a} \quad (3.2)$$

$\mathbf{V}$  は  $v(k, n)$  を要素とする  $K \times N$  の行列、 $\mathbf{a}$  は  $a_{rv}(k)$  を要素とする係数ベクトルである. 入力音声既知の場合、 $\mathbf{V}$  の逆行列を両辺にかけることで、係数ベクトル  $\mathbf{a}$  を求めることができる.

$$\mathbf{a} = \mathbf{V}^{-1}\mathbf{s} \quad (3.3)$$

ここで、雑音信号  $\mathbf{V}$  は  $N = K$  の正方行列でなければいけない.

このようにして求めた係数  $a_{rv}(k)$  を対応する雑音  $v(k, n)$  にかけることでサブ信号を

生成することができる。例えば、原音声を2つのサブ信号  $s_1, s_2$  に分解する場合、例えば以下のようにサブ信号を合成する。

$$\begin{cases} s_1(n) = \sum_{k=1}^{K/2} a_{rv}(k)v(k, n) \\ s_2(n) = \sum_{k=K/2+1}^K a_{rv}(k)v(k, n) \end{cases} \quad (3.4)$$

このとき、それぞれ加算する数は必ずしも  $K/2$  である必要はない。

ランダムベクトル分解法を用いて作成したサブ信号の性質は、あらかじめ用意した雑音信号に依存する。そのため、場合によっては生成したサブ信号の振幅が大きくなり、スピーカでの再生に適さない可能性がある。そこで、式(3.5)の様に、最小二乗法を用いることで少ない雑音信号で分解する手法が提案されている [52]。

$$\mathbf{a} = (\mathbf{V}^T \mathbf{V})^{-1} \mathbf{V}^T \mathbf{s} \quad (3.5)$$

これにより、用いる雑音信号数を少なくすることができるため、サブ信号のパワーを抑えることができる。しかし、最小二乗法を用いることで、式(3.1)を満たさず、すべての  $a_{rv}(k)v(k, n)$  を加算した再合成音声  $\hat{s}(n)$  に戻らない。そこで、原音声  $s$  から再合成音声を引いた誤差信号を求め、再合成音声に付加する。

ランダムベクトル分解法の問題点として、原信号を1サンプル単位で分解しているため、複数のサブ信号の位相がすべて揃わないと再合成音の音声品質が大きく低下する。したがって、再生制御点上ではもとの信号を復元できたとしても、その周囲では音声品質が大きく低下することが考えられる。また、1フレームのデータ点数よりも極端に少ない雑音信号数によって分解した場合、最小二乗法による誤差が大きくなり、結果として誤差信号中の原信号成分が多くなる可能性がある。

### 3.3.2 ノイズ駆動分解法

ノイズ駆動は音声をLPC分析し、得られた残差信号と任意の雑音信号との差分信号をLPC合成することでサブ信号を得る手法である。LPC分析・合成を用いることで、

合成時に音声の特徴が復元されるため、ある程度音声の特徴を残した分解ができると考えられる。

図 3.2 にノイズ駆動分解法の流れを示す。初めに、入力信号に対して LPC 分析を適用する。LPC 分析は時刻  $n$  の入力サンプルを LPC 係数と過去サンプルの値から予測する線形予測モデルであり、主に音声符号化や音声解析などで用いられる。入力信号  $x(n)$  の予測値を  $\hat{x}(n)$  とすると、以下のように表すことができる。

$$x(n) = \hat{x}(n) + e(n) = - \sum_{i=1}^p a_{\text{lpc}}(i)x(n-i) + e(n) \quad (3.6)$$

ここで、 $p$  は LPC 次数、 $a_{\text{lpc}}(i)$  は LPC 係数、 $e(n)$  は残差信号である。 $a_{\text{lpc}}(i)$  は  $e(n)$  の二乗誤差が最小になるように求める。

本論文では残差信号を異なるガウス雑音を用いて複数の信号に分解する。ガウス雑音とはガウス性を満たす白色雑音である。例えば 2 つのサブ信号を生成する場合、異なるガウス雑音  $z_1(n)$ ,  $z_2(n)$  を用いて残差信号との差分  $y_1(n)$ ,  $y_2(n)$  を計算する。

$$\begin{cases} y_1(n) = e(n) - z_1(n) \\ y_2(n) = e(n) - z_2(n) \end{cases} \quad (3.7)$$

求めた差分信号を相互に加算することで、サブ信号の残差信号  $e'_1(n)$ ,  $e'_2(n)$  を得る。

$$\begin{cases} e'_1(n) = z_1(n) + y_2(n) \\ e'_2(n) = z_2(n) + y_1(n) \end{cases} \quad (3.8)$$

この残差信号  $e'_1(n)$ ,  $e'_2(n)$  を式 (3.6) の  $a_{\text{lpc}}(i)$  を用いてそれぞれ LPC 合成を行うことでサブ信号を得る。LPC 合成フィルタは LPC 分析フィルタの逆フィルタとなっている。

このように、任意の雑音信号に対して、もう一方の雑音信号と残差信号の差分信号を加算することで差分信号の残差信号を得る。この差分信号の残差信号は足し合わせることで原信号の残差信号と一致する。LPC 合成フィルタは線形フィルタであるため、それぞれの信号を LPC 合成し生成したサブ信号を加算することで、原信号に復元できる。

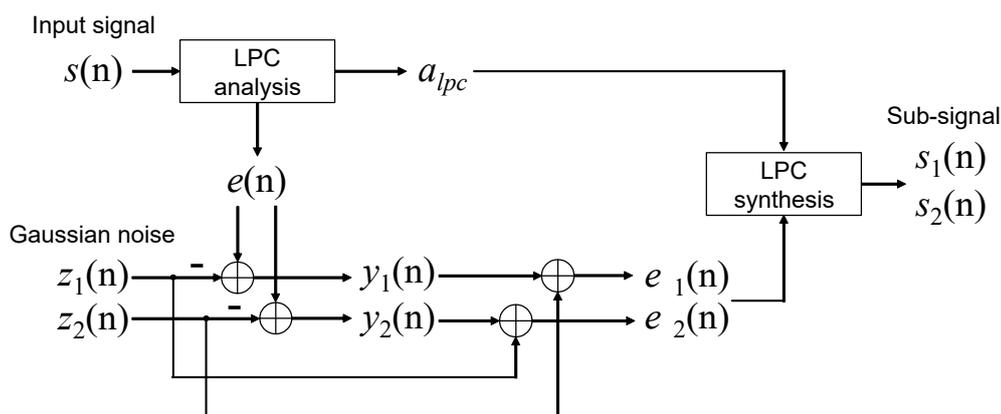


図 3.2: ノイズ駆動分解法の処理の流れ

### 3.3.3 残差ランダムベクトル分解

ランダムベクトル分解法はサブ信号のランダム性が強く、目的方向での再生品質に問題が生じる可能性がある。一方、LPC 残差信号を用いたノイズ駆動分解法は、サブ信号に音声成分が多く残り、目的方向外での音声劣化が少ない可能性がある。そこで、ランダムベクトル分解法と LPC 分析を組み合わせた手法を検討する。

図 3.3 にサブ信号生成の流れを示す。初めに入力信号を LPC 分析し、LPC 係数と残差信号に分解する。得られた残差信号に対して、ランダムベクトル分解を行い、2つの信号に分解する。この信号をそれぞれ LPC 合成することで、2つのサブ信号を得る。ランダムベクトル分解には最小二乗法を用いるが、前述のように分解に用いる雑音信号数によっては分解後の信号が安定しない可能性がある。

### 3.3.4 雑音付加によるサブ信号の安定化

前項の分解法では、残差信号をランダムベクトル分解によって分解するため、場合によっては LPC 合成に適さない信号が入力され、サブ信号が安定しない可能性がある。そこで、ランダムベクトル分解を行った残差信号にガウス雑音を付加し、LPC 合成フィルタに入力する信号の安定化を図る。図 3.4 に処理の流れを示す。ガウス雑音を付加する際には、ランダムベクトル分解した残差信号とガウス雑音の比率を調節できるよう

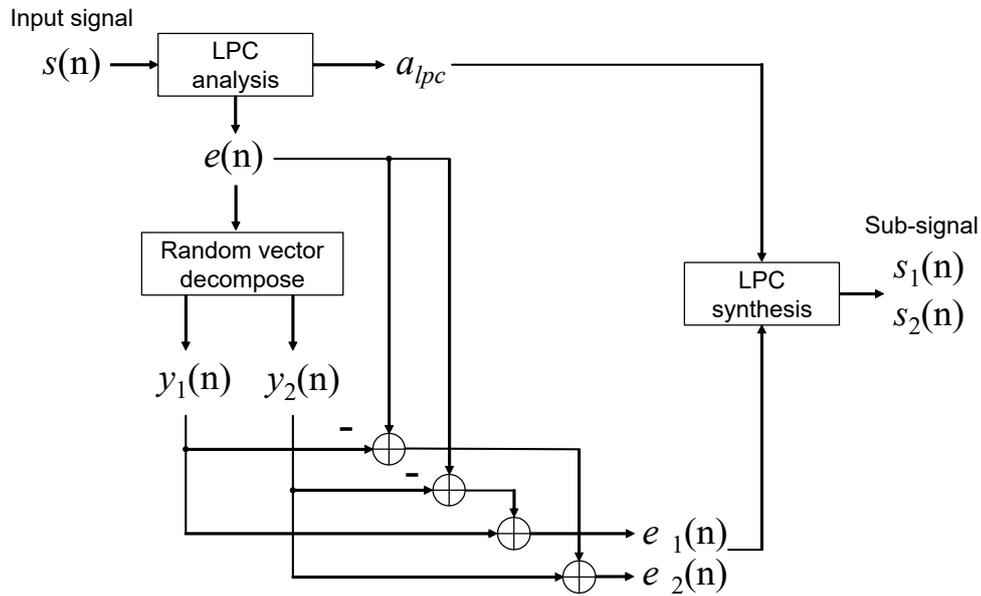


図 3.3: 残差信号へのランダムベクトル分解の適用

に比率係数  $a(0 \leq a \leq 1)$  を導入し、以下のように加算する。

$$\begin{aligned}
 y_1 &= ax'_1 + (1-a)z_1 \\
 y_2 &= ax'_2 + (1-a)z_2
 \end{aligned}
 \tag{3.9}$$

比率係数  $a$  を小さくするほどガウス雑音の割合が大きくなるため、LPC 合成後の信号が安定する。しかし、サブ信号同士の特徴が近くなるため、再生時に目的方向外での音声の劣化が小さくなると考えられる。また、最後に差分信号を求め、相互に加算することで、再合成時の原信号への復元を保証する。

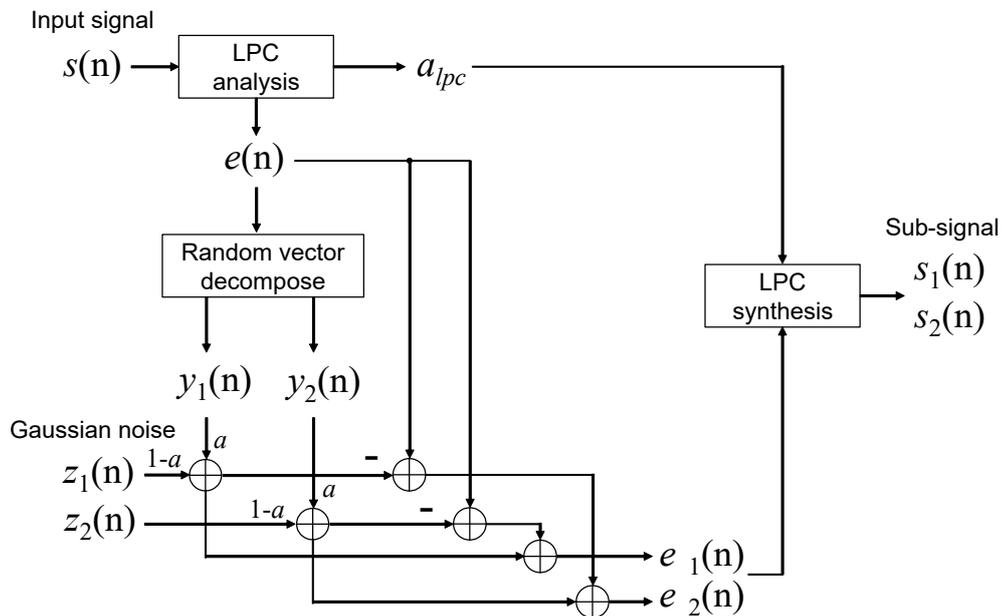


図 3.4: ガウス雑音の付加によるサブ信号の安定化

## 3.4 評価実験

信号分解法と多点制御法を用いたエリア再生手法について、相対音圧レベルによる客観評価実験と DMOS 評価法による主観評価実験を行った。本節では、それぞれの実験条件と結果について述べる。

### 3.4.1 客観評価実験

実験に使用した信号は男女の各 5 名計 10 名の発話音声である。サブ信号数は 2 とし、図 3.1 のように異なるスピーカアレーから再生した。スピーカ数は各アレー 4 個、合計 8 個である。ランダムベクトル分解法のフレーム長は 32 とし、使用した雑音数は 16 とした。また、ノイズ駆動分解法、残差ランダムベクトル分解法で用いる LPC 分析・合成の次数は 32、フィルタ長は 512 とした。サンプリングレートは 16 kHz である。

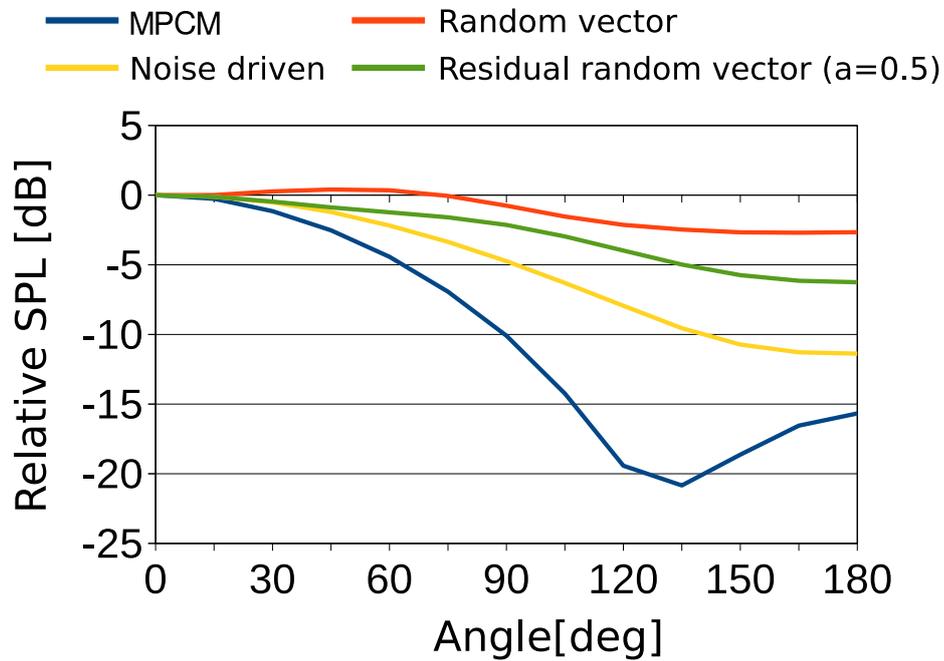


図 3.5: 各信号分解手法の相対音圧レベル

図 3.5 に各手法の目的方向を基準とした相対音圧レベルを示す。図中の MPCM は 8 個のスピーカで多点制御法を行った場合の結果である。残差ランダムベクトル分解法の付加比率は  $a = 0.5$  である。図 3.5 から、信号分解法を用いた場合、多点制御法と比べて目的方向外での音圧が上昇している。これは、信号分解法により作成したサブ信号を多点制御法によって再生するため、スピーカアレーを分割したことで、スピーカ数が少なくなり指向性が弱くなったためである。また、それぞれの分解法では、ランダムベクトル分解法、残差ランダムベクトル分解法、ノイズ駆動分解法の順に目的方向外での抑圧量が大きくなっている。これは、多点制御法が主に低域に強い手法であるため、サブ信号に含まれる帯域が音声に近いほど目的方向外での抑圧量が大きく、白色雑音に近いほど抑圧量が小さくなっていると考えられる。

### 3.4.2 主観評価実験

評価カテゴリには表 2.2 の劣化についての評価語を用いた。図 3.6 に実験結果を示す。実験条件は音声サンプル数は男女各 2 名の発話音声を計 10 音声，被験者数は男女計 12 名である。音声は計算機シミュレーションによりそれぞれの位置での音声を作成し，ヘッドホンにより受聴した。各手法の条件は客観評価試験と同様である。

図 3.6 から，多点制御法の再生誤差によってサブ信号が完全に復元できていないため，ランダムベクトル分解法では目的方向での劣化が大きく DMOS 値が小さくなっている。ノイズ駆動分解法は目的方向での音質を維持できており，品質に問題がないことが確認できる。また，目的方向外では，ランダムベクトル分解法が最も劣化していることがわかる。残差信号分解法，残差ランダムベクトル分解法は，ランダムベクトル分解法と比べると品質劣化は少ないが，多点制御法と比べると大幅に劣化しており，聞き取り難くなっていると考えられる。

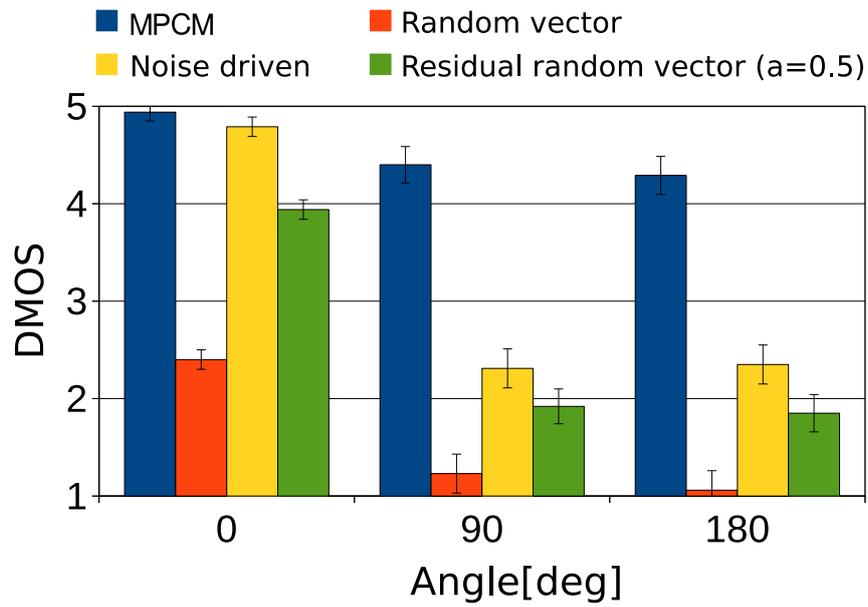


図 3.6: 各信号分解手法の DMOS 評価結果

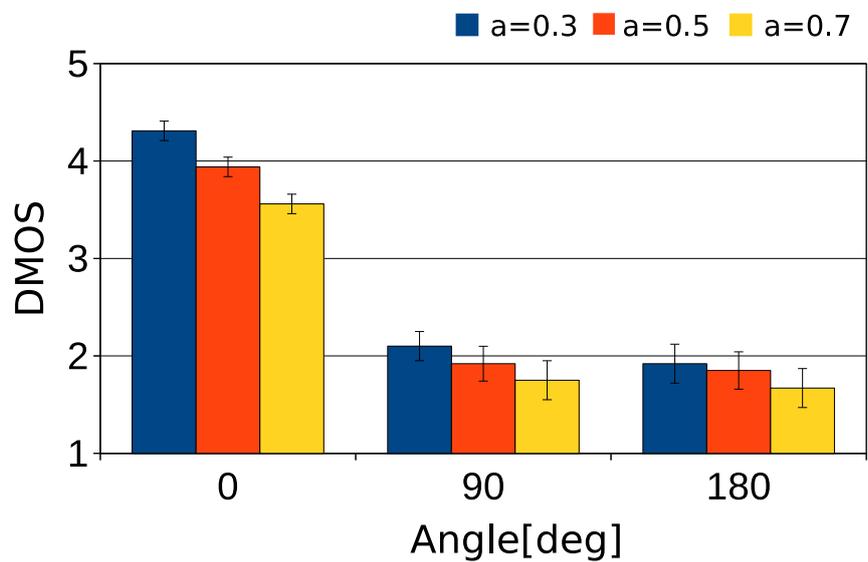


図 3.7: ガウス雑音の比率係数の違いによる DMOS 値の変化

## 3.5 まとめ

本章では，多点制御法と信号分解法を組み合わせた用いたエリア再生手法を提案した．また，音声をランダムに分解する手法とLPC分析を用いて，音声の特徴を残して分解する手法を提案した．それぞれの分解手法を用いたエリア再生手法について，客観評価実験と主観評価実験を行った．客観評価結果から，従来の多点制御法と比べて目的方向外での抑圧量は低下したが，主観評価結果から目的方向外での音声の劣化量は大きくなっており，聞き取り難さが上昇していることが確認できた．

信号分解法を用いたエリア再生手法は，サブ信号を再生するためにそれぞれ独立したアレーを用意する必要がある．したがって，1つの入力信号に使用できるスピーカ数が減少するため，目的方向外での抑圧量が低下する．また，サブ信号の性質は入力信号と用意する雑音信号によって決まるため，場合によっては予期せず不快感を増大させる可能性がある．そのため，第4章において入力信号によらずに，目的方向外の聞き取り難さを上昇させる手法を検討する．

# 第4章 目的方向外の音声に対する 環境音によるマスキング手法

## 4.1 はじめに

多点制御法を用いたエリア再生において、目的方向外での音声劣化が少なく漏えいした音声内容が聞き取れる問題があった。前章では信号分解法を用いて再生する音声を分解し、空間上で再合成する手法を提案したが、再生に用いるスピーカ数が減少するため、多点制御法の性能が低下する問題があった。本章では、マスキングによるスピーチプライバシー保護の手法を導入し、目的方向外に環境音を再生することで、漏えいした音声を聞き取り難くする手法を提案する。

以下、4.2節において、スピーチプライバシー保護のためのマスキング手法の種類について述べる。4.3節において、多点制御法を用いた環境音の制御手法について述べる。4.4節において、客観評価について述べる。4.5節において、主観評価実験について述べる。4.6節において、本章のまとめを述べる。

## 4.2 スピーチプライバシー保護のためのマスキング手法

会話内容を保護するスピーチプライバシーという考え方がある [53, 54]。これまでにスピーチプライバシーを保護する様々な手法が提案されているが、電気音響的には会話音声に対してスピーカから別の音を再生することで聞き取れなくする方法が用いられている [55, 56, 57]。

スピーチプライバシー (Speech Privacy) という用語の意味は日本と欧米でわずかに異なっている [53, 58]。国内では個人の私事に関する会話の漏えいに対する評価語とし

て捉えられているが、欧米では作業環境が他の環境と切り離され、集中できる環境になっているかを評価する用語である。したがって、会話の漏えいが防ぐことができているかどうかだけでなく、その環境に対する評価も含まれている。会話の漏えいのみを評価する用語としてはスピーチセキュリティ(Speech Security)が用いられている。本研究では、音声内容の聞き取りに焦点を当てるため、スピーチプライバシーを国内で認識されている、個人の私事に関する会話の漏えいに対する評価語として考える。

スピーチプライバシーを保護する手法としてABCルール(Absorb: 吸音, Block: 遮音, Cover-up: マスキング)が知られている。AbsorbとBlockについては空間を設計する段階で考慮する必要があり、既存の施設では導入が難しい場合がある。Cover-upについてはスピーカが設置できる環境であれば容易に導入することができる。

マスキングとは、ある可聴音を聞き取るための最小可聴値が他の音によって上昇し、聞き取りにくくなる現象である。マスキングされる音をマスク(Maskee)、マスキングに用いる音をマスク(Masker)という。本研究では、マスクとして定常雑音や会話音声などの環境音を使用し、多点制御法によって制御を行う。

#### 4.2.1 パワーマスキング

パワーマスキングはあるマスクに対して、マスクをある程度の大きさで同時に再生することによってSN比を低下させ、聞き取りを妨害する手法である。一般的には音楽BGMや流水音など不快感を伴わない音が用いられる。定常雑音を用いる場合は帯域制限をしたピンクノイズが最も有効であることが知られている[55]。パワーマスキングによるスピーチプライバシー保護手法は、もっとも単純で簡単な手法であるが、状況に応じたマスクの選択が難しいことや、マスクの特徴や大きさなどによっては、十分なマスキング効果が得られず、不快感を与えることがある。

## 4.2.2 情報マスキング

音声内容や話者が既知の場合，話者本人の音声特性を持つ信号のマスキング効果が高いことが報告されている [59]．このように音声の特徴を用いて効率よくマスキングする手法を情報マスキングと呼ぶ．情報マスキングはパワーマスキングと比べ，比較的小さい音でマスキングできるため，受聴者に対して不快感を与えにくい手法である．しかし，事前に話者本人の音声を取得することは難しく，信号処理が必要な場合，マスカが再生されるまでの遅延を考慮する必要がある．

## 4.3 多点制御法による環境音の指向性制御

多点制御法において目的方向外へ漏えいした音声のマスキングを行う．ここで，環境音についても多点制御法を用いて目的方向外に向けて再生を行い，目的方向への環境音の漏えいを減少させる．本研究では， $30 \sim 180 \text{ deg}$  を目的方向外とするため，図 4.1 のように  $30 \sim 180 \text{ deg}$  を再生制御点としたフィルタを設計する．このフィルタ構成を一括方式と呼ぶ．多点制御法を用いることで目的方向外のみ環境音を再生するが，目的音声と同様に環境音が目的方向へ漏えいすることが考えられる．

図 4.2 に  $30 \text{ deg} \sim 180 \text{ deg}$  を再生制御点とした場合の  $180 \text{ deg}$  を基準とした相対音圧レベルを示す．マスカとして使用した環境音はホワイトノイズ，ピンクノイズ，Multiple Talkers である．Multiple Talkers は目的音声とは異なる話者の音声をランダムな区間で複数重畳したものであり，人の声であることは認識できるが，発声内容は理解できない．図 4.2 から， $0 \text{ deg}$  での音圧レベルが  $-3 \sim -8 \text{ dB}$  となっており，目的方向において環境音がわずかに漏えいしていることがわかる．

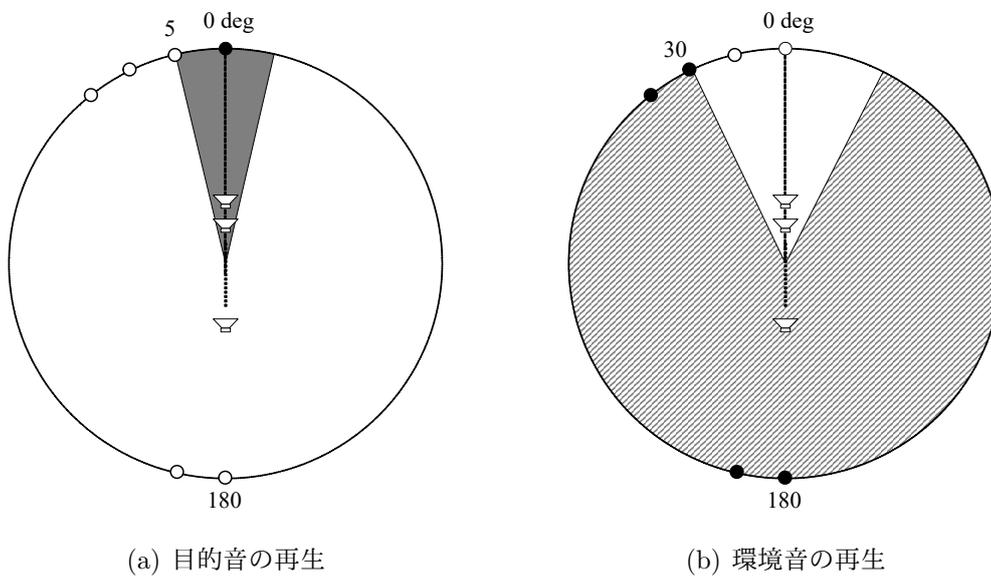


図 4.1: マスキング手法を用いた多点制御法の制御点配置

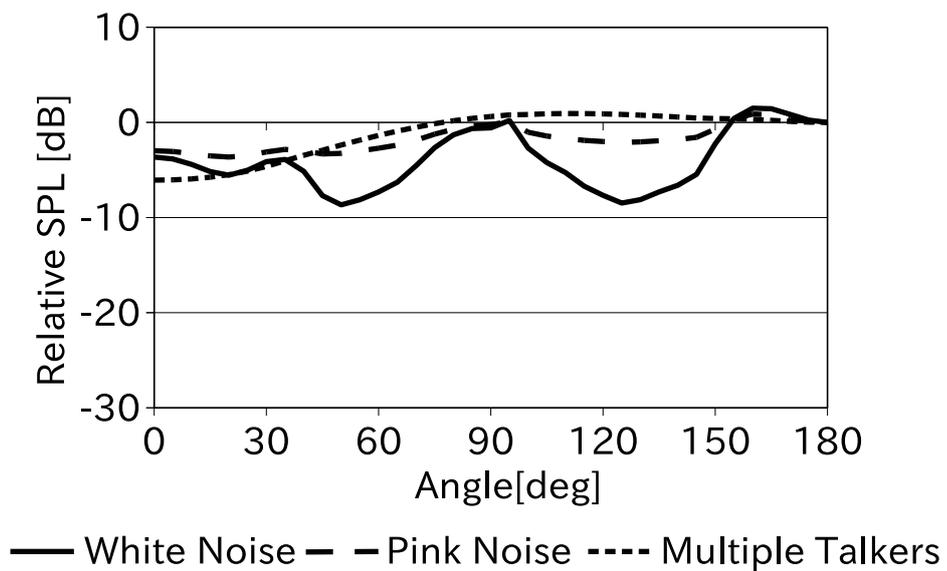


図 4.2: 一括方式で再生した環境音の相対音圧レベル

### 4.3.1 逆位相音の再生

目的方向への環境音の漏えいを防ぐために環境音の逆位相信号を目的方向へ再生する。多点制御法によって目的方向へ環境音の逆位相信号を再生することで、目的方向では環境音が打ち消され、目的音声を十分な音質で再生できる。

環境音にとって目的方向外となる0 degでは周波数特性が乱れる。周波数特性が乱れると目的方向へ漏えいする環境音の波形が変形するため、逆位相信号の効果を得られない。そこで、式(2.8)の拘束条件付き多点制御法を用いて周波数特性の補正を行い、フィルタ係数を求める。

図4.3に一括方式を用いた場合の拘束条件による周波数特性の変化を示す。図4.3から、拘束条件を0 degの制御点にすることで周波数特性が改善していることがわかる。

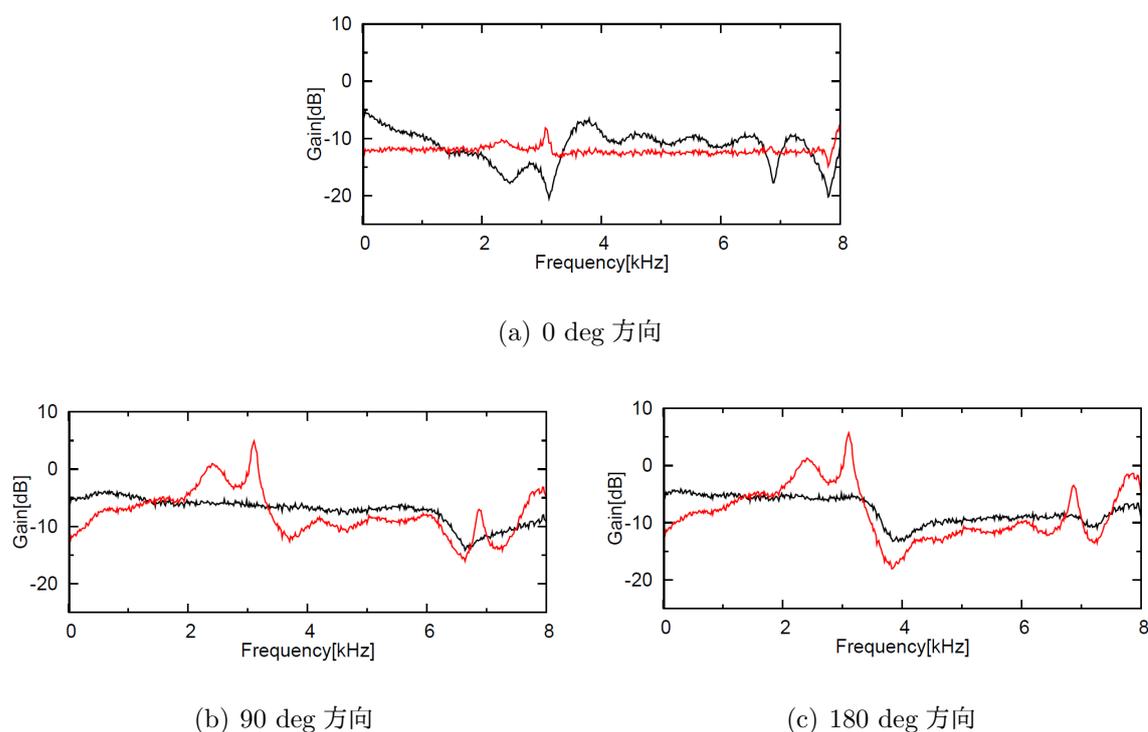


図 4.3: 各方向の周波数特性 (黒：拘束条件変更前 赤：拘束条件変更後)

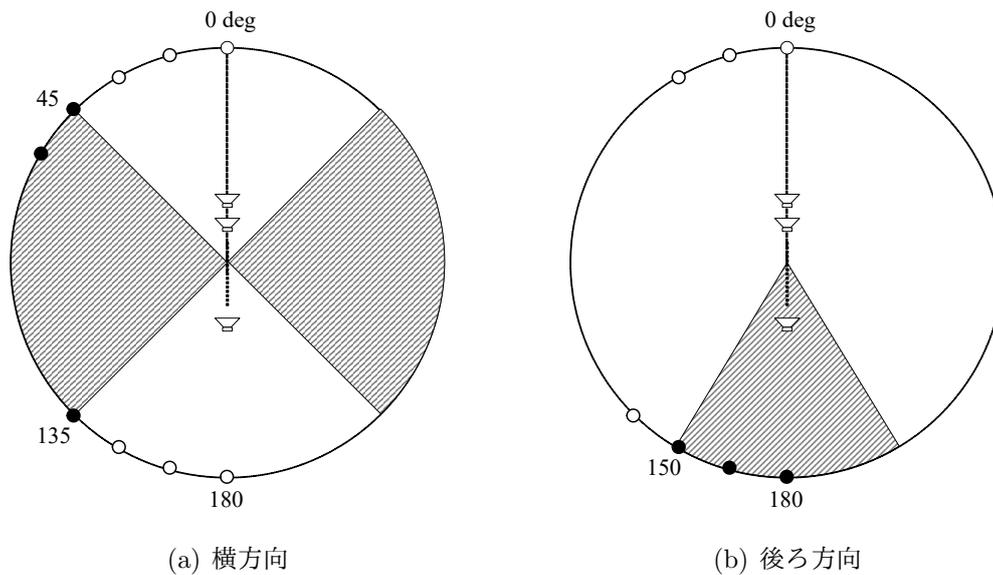


図 4.4: 環境音を分割再生するときの制御点配置

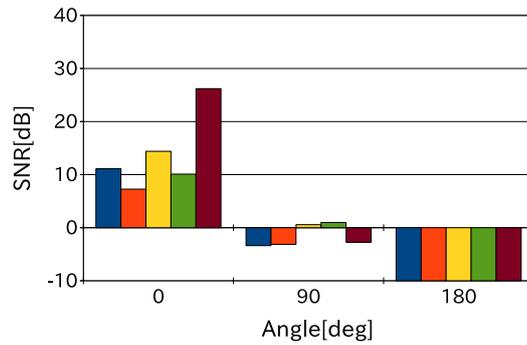
### 4.3.2 環境音再生フィルタの分割

目的方向への環境音の漏えいを抑えるため、環境音を再生するフィルタを2つに分割する。図 4.4 に制御点配置を示す。分割する範囲については予備検討から 45 ~ 135 deg へ再生する Filter A, 150 ~ 180 deg へ再生する Filter B とした。環境音の指向性を高め、30 ~ 45 deg, 135 ~ 150 deg に再生を行わないことで、目的方向への漏えいを少なくする。この方式を分割方式と呼ぶ。

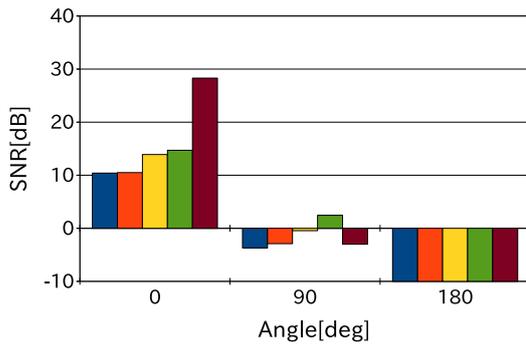
## 4.4 客観評価実験

これまでに述べた再生方式について、180 deg を基準とした相対音圧レベルを図 4.5 に示す。図 4.5(a) から、分割方式では Multiple Talkers が 0 deg で約-21 dB と抑圧できているが、ホワイトノイズ、ピンクノイズは約-8 dB となっており、目的方向への漏えいが改善できていない。また、図 4.5(b)(c) の逆位相信号を付加する方式では、逆位相信号を付加しない方式と比べ、目的方向での抑圧量が増加しておらず、効果が見られない。図 4.5(d) の拘束条件を変更し、逆位相信号を付加する方式は、目的方向での抑圧量がホワイトノイズとピンクノイズで約-20 dB, Multiple Talkers で約-26 dB となっ

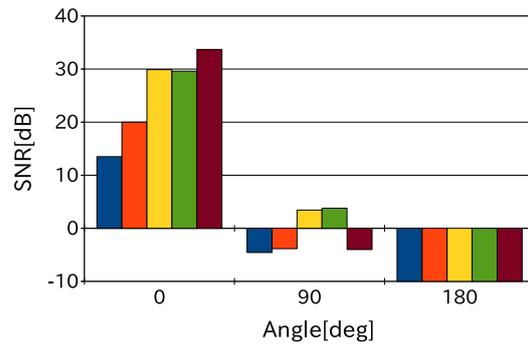




(a) ホワイトノイズ



(b) ピンクノイズ



(c) Multiple Talkers

図 4.6: 180 deg において目的音声に対して-10 dB としたときの各方向の SNR

dB となっている。しかし、他の方式と比べ、分割方式は 90 deg 方向での SNR が大きくなっていることから、目的方向外、特にアレーに対して横方向でのマスキング効果が低いと考えられる。また、一括方式の拘束条件を変更し逆位相信号を適用したものが目的方向において、すべての環境音で 25 dB を超えている。したがって、環境音によらず、目的方向での聞き取りに影響はないと考えられる。

## 4.5 主観評価実験

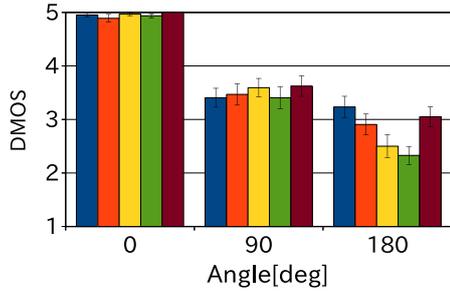
### 4.5.1 実験条件

フィルタ構成の違いによる目的方向外での聞き取り難さと不快さを評価するため、DMOS 評価法による主観評価実験を行った。音声サンプル数は男女各 2 音声。環境音の音圧は 180 deg において SNR が 0, -5, -10 dB とした。被験者は 20 代の男女計 16 名である。評価には表 2.3 の聞き取り難さと不快さについての評価語を用いた。

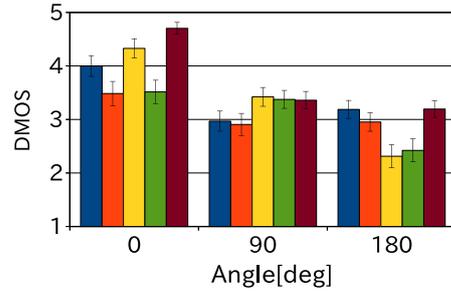
### 4.5.2 実験結果

図 4.7-4.9 に DMOS 評価の結果を示す。結果から、従来の一括方式は目的方向での不快さが大きくなっている。これは、目的方向で環境音が十分に抑圧されず、漏えいしていることが影響していると考えられる。また、フィルタを分割する方式と拘束条件を変更した一括方式に逆位相を付加したものが目的方向での不快さが少ない。フィルタを分割する方式では、分割したことによって目的方向への漏えいが少なくなったためであり、拘束条件を変更した一括方式は逆位相信号の効果によるものと考えられる。フィルタを分割する方式は 90 deg での聞き取りにくさが高くなっており、聞き取り易くなっている。

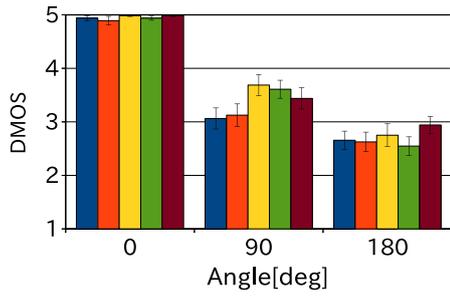
目的方向での不快さを考慮し、目的方向外での聞き取り難さを考えると、拘束条件を変更した一括方式に逆位相信号を付加したものが最もよい結果であるといえる。



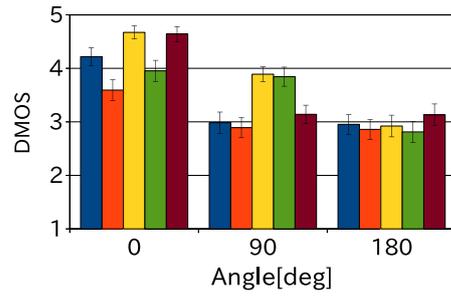
(a) ホワイトノイズによる聞き取り難さ



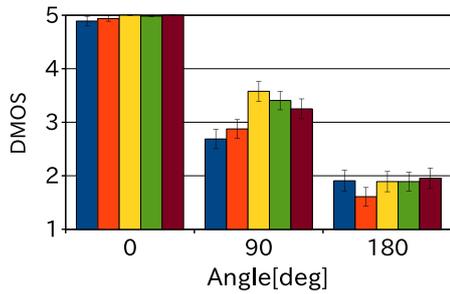
(b) ホワイトノイズによる不快さ



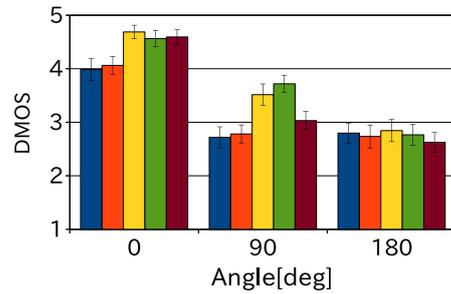
(c) ピンクノイズによる聞き取り難さ



(d) ピンクノイズによる不快さ

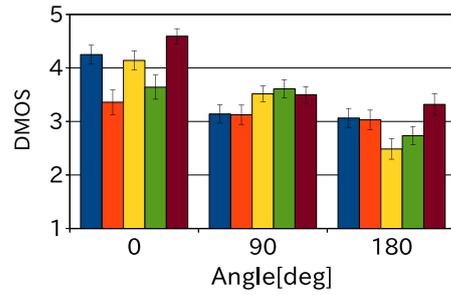
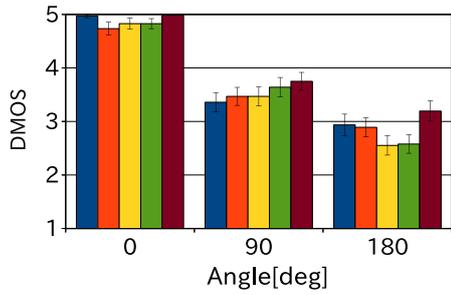


(e) Multiple Talkers による聞き取り難さ



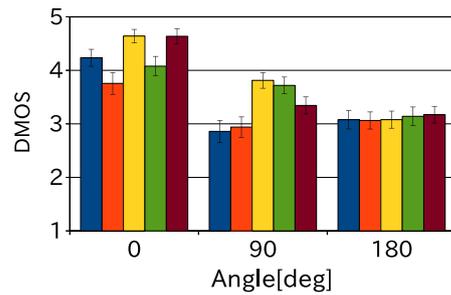
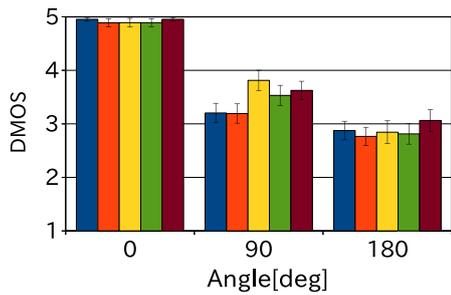
(f) Multiple Talkers による不快さ

図 4.7: DMOS 評価による主観評価実験結果 ( $SNR_{180} = -10$ )



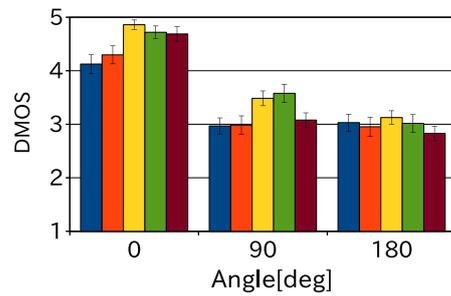
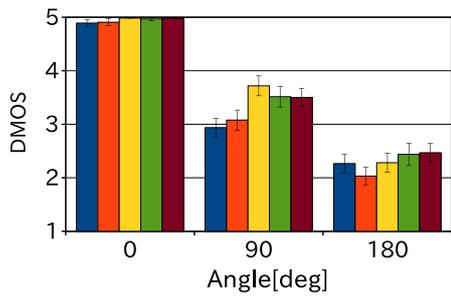
(a) ホワイトノイズによる聞き取り難さ

(b) ホワイトノイズによる不快さ



(c) ピンクノイズによる聞き取り難さ

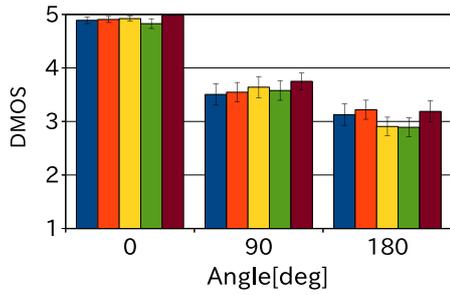
(d) ピンクノイズによる不快さ



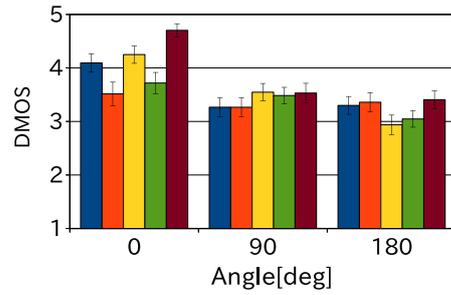
(e) Multiple Talkers による聞き取り難さ

(f) Multiple Talkers による不快さ

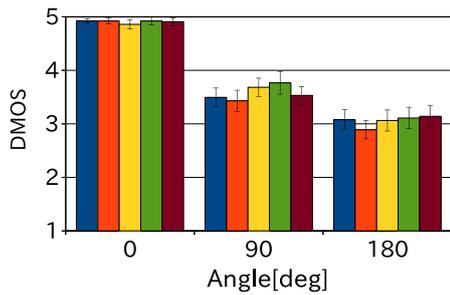
図 4.8: DMOS 評価による主観評価実験結果 ( $SNR_{180} = -5$ )



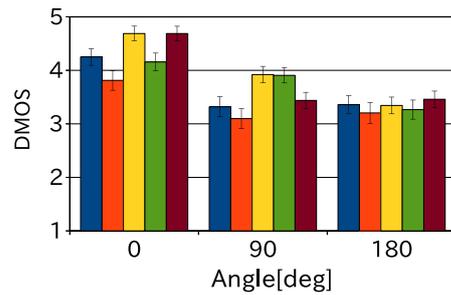
(a) ホワイトノイズによる聞き取り難さ



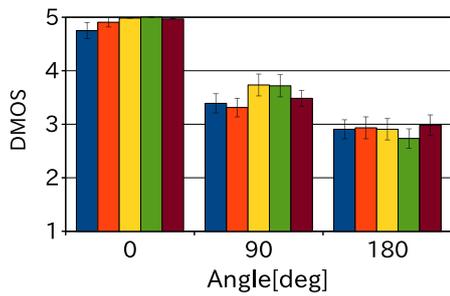
(b) ホワイトノイズによる不快さ



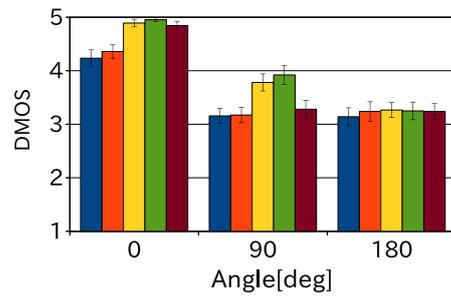
(c) ピンクノイズによる聞き取り難さ



(d) ピンクノイズによる不快さ



(e) Multiple Talkers による聞き取り難さ



(f) Multiple Talkers による不快さ

図 4.9: DMOS 評価による主観評価実験結果 ( $SNR_{180} = 0$ )

## 4.6 まとめ

本章では，多点制御法によるエリア再生手法において，目的方向外へ漏えいする音声に対して環境音を再生する手法を提案した．環境音を再生するときにも多点制御法による制御を行い，フィルタを分割して再生する手法を提案した．また，目的方向へ環境音の逆位相信号の再生を行い，目的方向への環境音の漏えいを防ぐ手法を提案した．

客観評価実験から，目的方向に環境音の逆位相信号を再生することで，環境音を目的方向外へ均一に再生しつつ，目的方向での SNR を改善できることを確認した．また，主観評価実験から，環境音として Multiple Talkers を用いた場合に，目的方向外で最も聞き取り難くなることを確認した．

今回の実験では計算機シミュレーションによって作成した音声を評価しているが，実際の環境では部屋の環境やスピーカの設置誤差によって，逆位相信号による効果が期待できない可能性がある．したがって，リファレンス用のマイクなどを設置し，適応的に環境音を再生できるような手法を検討する必要がある．



# 第5章 複数エリア再生

## 5.1 はじめに

ある程度広い範囲への音のエリア再生は博物館や美術館での展示案内音声や駅での自動放送案内システムなど、非常に有用な技術である。さらに、再生するエリアを2つ以上設定する複数エリア再生の技術は多言語での案内放送やデジタルサイネージなど幅広い応用が期待できる。本章では、多点制御法を用いて複数のエリアに再生する手法について述べる。また、より少ない抑圧制御点で実現するための制御点配置の検討や実環境での評価実験について述べる。

以下、5.2節において、多点制御法を用いて複数のエリアに異なる音声を再生する方法について述べる。5.3節において、複数エリア再生を行う際の制御点配置について述べる。5.4節において、計算機シミュレーションによる客観評価試験について述べる。5.5節において、実環境での客観評価と単語了解度試験について述べる。5.6節において、本章のまとめを述べる。

## 5.2 複数エリア再生

多点制御法はスピーカから制御点までの伝達関数に基づいて、音圧特性を制御するフィルタを設計する。そのため、スピーカの配置や制御点配置に制約が少なく、比較的自由度の高い配置が可能である。しかし、実環境で制御するためには、インパルス応答を測定する必要があるため、測定が容易な制御点配置であることが望ましい。多点制御法を用いて複数のエリアに異なる音声を再生するためには、それぞれのエリアに再生するフィルタを設計し、それぞれの出力を重ねることで実現が可能である。本研

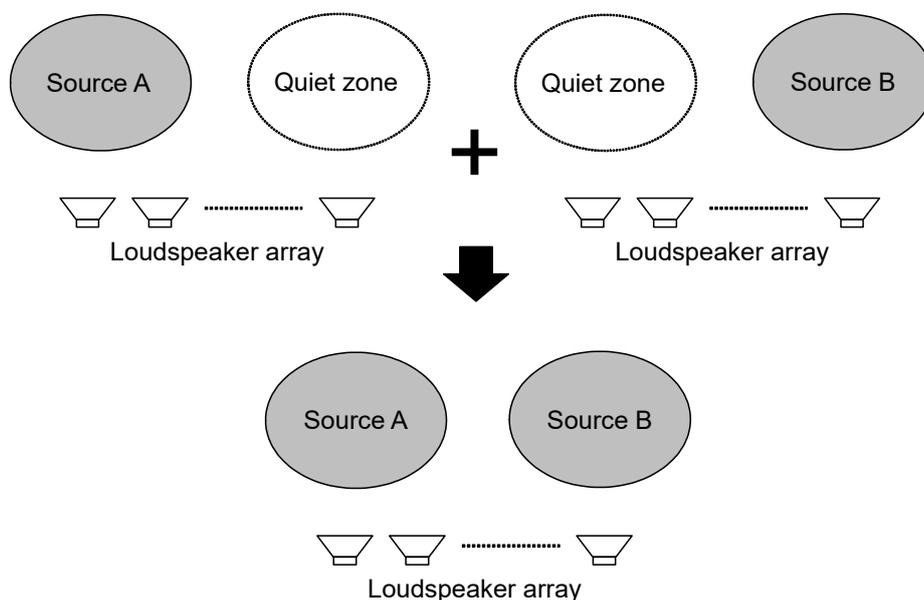


図 5.1: 複数エリア再生のイメージ

究では目的音声を再生するエリアを再生エリア，抑圧するエリアを抑圧エリアと呼ぶ．図 5.1 に複数エリア再生のイメージ図を示す．図 5.1 中の左上の図はアレーの左側を再生エリア，右側を抑圧エリアとし，左側への再生を行うことができる．同様に，右側へ再生するフィルタを作成し，それぞれの出力を足し合わせることで，それぞれのエリアに異なる音声を再生することができる．本論文では再生点でのみ音が再生されるのではなく，ある程度広いエリアでの再生を目指す．また，直線アレーを用いるため，制御はスピーカアレーと同一平面上のみを考える．

### 5.3 制御点配置の検討

図 5.2 に本研究で用いるスピーカアレーとエリアの配置図，再生点の位置を示す．スピーカアレーに平行な方向を  $x$ ，垂直な方向を  $y$  とし，スピーカアレーの中心を  $x = 0$ ， $y = 0$  とする．再生を行う Area A の再生制御点はスピーカアレーの中央から左側に 1 m，前方に 1 m の点 ( $x = -1.0$ ， $y = 1.0$ ) とした．

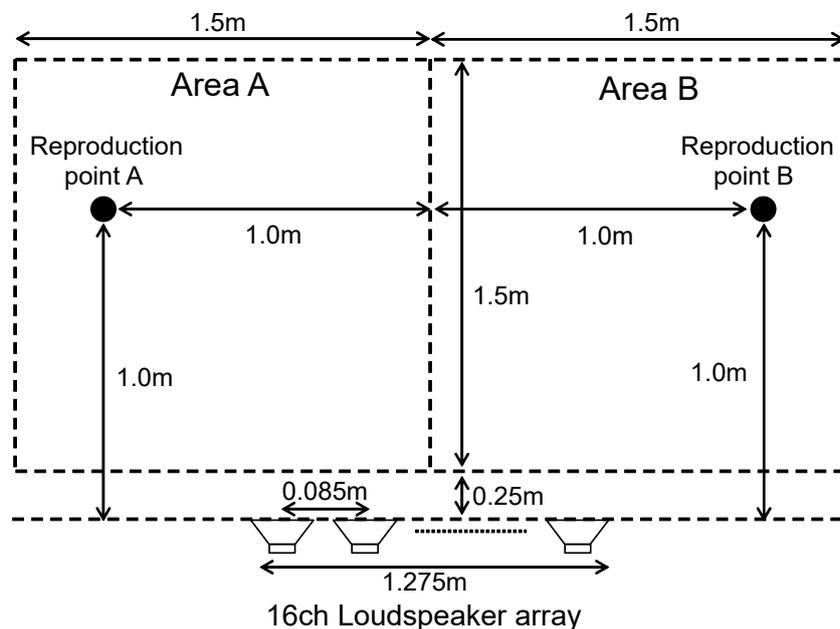


図 5.2: 制御するエリアと再生制御点の配置

通常、多点制御法では再生制御点の周囲に抑圧制御点を設置し、指向性を制御するが、複数エリア再生においては Area A 全体に再生音が一樣に再生されることが望ましい。そこで、Area B 内に抑圧制御点を配置する方法を検討する。できるだけ少ない制御点数で再生エリアでの音質を保ち、抑圧エリアでの音圧を小さくするため、3種類の制御点配置について検討を行った。

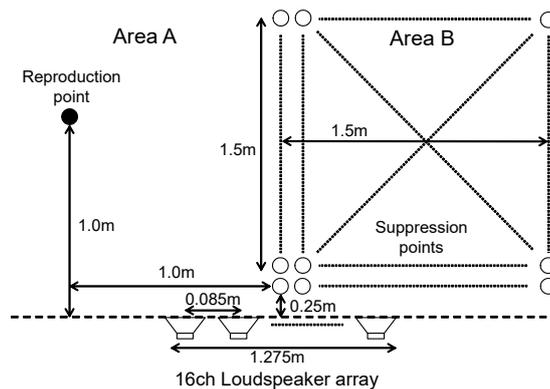
図 5.3 に具体的な制御点配置を示す。まず、図 5.3(a) のように、Area B の内部に抑圧制御点を敷き詰める配置が Area B 全体を最も抑圧できると考えられる。しかし、制御点数が膨大になると再生制御点での再生品質への影響が考えられる。また、実環境ではインパルス応答の測定が必要であり、測定時間はスピーカ数と制御点数の組み合わせに、測定信号長と測定装置の設置時間をかけたものとなるため、制御点数は少ないほうが良い。そこで、図 5.3(b) のように、Area B の周囲のみに抑圧制御点を配置する。また、図 5.3(c) のように、スピーカアレーから遠い 2 辺に関しては、音波の進行方向を考えると Area B の外側となるため、スピーカアレーに近い 2 辺のみの配置につ

表 5.1: 各配置の制御点数 (再生制御点を含む)

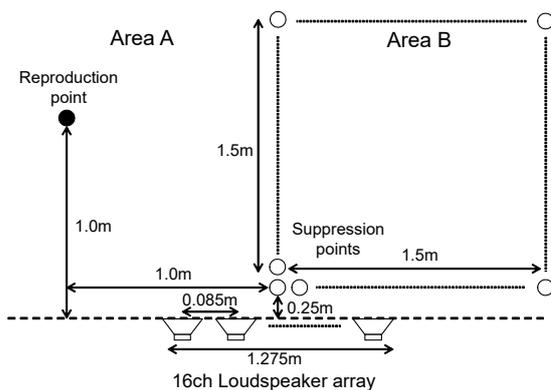
抑圧制御点間隔	格子配置	境界配置	L型配置
0.1 m	257	63	32
0.02 m	5777	303	152

いて検討する。

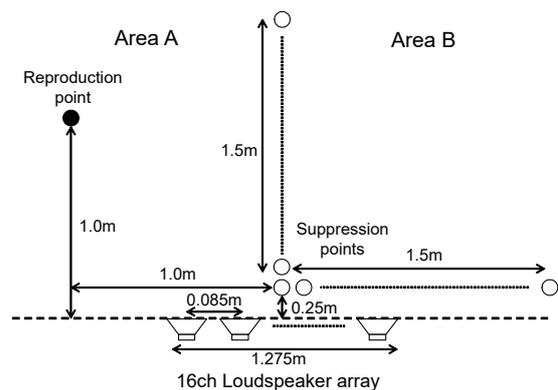
抑圧制御点の間隔については、間隔が狭いほどエリア全体の抑圧量は向上すると考えられるが、式 (2.5) の空間相関行列の独立性が低くなり、フィルタが不安定になりやすく、制御点数の増加にもつながる。そこで抑圧制御点間隔 0.1 m と 0.02 m の 2 種類について実験を行った。表 5.1 に再生制御点 1 点を含む、各配置の制御点数を示す。表より、格子配置は制御点数が非常に多いことがわかる。



(a) 格子状配置



(b) 境界上配置



(c) L型配置

図 5.3: Area A に再生を行うときの制御点の配置

## 5.4 計算機シミュレーション

表 5.1 の 6 種類の制御点配置について計算機シミュレーションによる評価実験を行った。

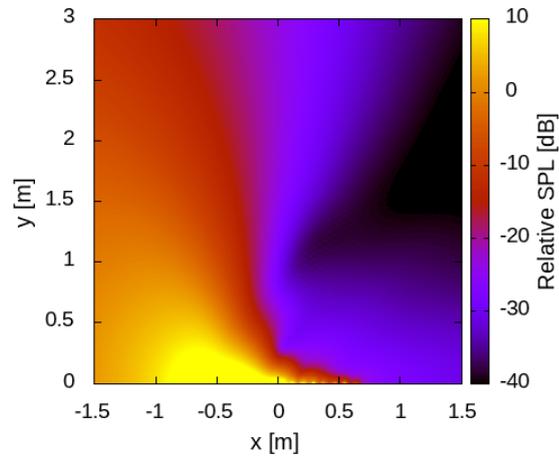
### 5.4.1 実験条件

シミュレーションは自由空間を仮定し，伝達関数には 3 次元グリーン関数を用いた。スピーカ数は 16ch，スピーカ間隔は 0.085 m とした。評価指標として，再生点の音圧を 0 dB としたときの相対音圧と PESQ(Perceptual Evaluation of Speech Quality)[60]を用いた。PESQ は ITU-T P.862 で定義される音声品質の客観評価法であり，知覚モデルと認知モデルから MOS の推定値を出力する。したがって，評価値 (PESQ 値) は 5 に近いほど音声品質が高く，1 に近いほど音声品質が悪いことを意味する。使用した音声 (300 Hz~3.4 kHz) は男女各 4 名，計 8 名の内容の異なる 8 秒間の文章各 12 個，計 96 個である。

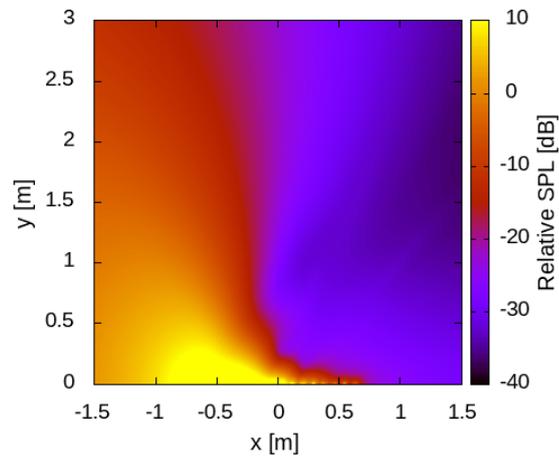
### 5.4.2 実験結果

図 5.4 と図 5.5 に相対音圧マップの結果を示す。音圧マップはすべての周波数の平均値である。色が暗いほど音が小さく，明るいほど音が大きいことを表している。図中の  $x = 0$  m,  $y = 0$  m がスピーカアレーの中心であり，中心から左側を再生エリアとしている。図から，どの制御点配置でもスピーカアレー右側の抑圧エリアである程度の抑圧ができていることがわかる。最も制御点の少ない図 5.4(c) の制御点間隔 0.1 m の L 型配置では抑圧制御点の隙間から抑圧エリアへ音の漏えいが見られる。また，図 5.4(d) の格子配置の 0.02 m 間隔では抑圧量が高く，逆に， $x = 0$  m 付近の再生エリアにおいて抑圧されてしまっていることが確認できる。

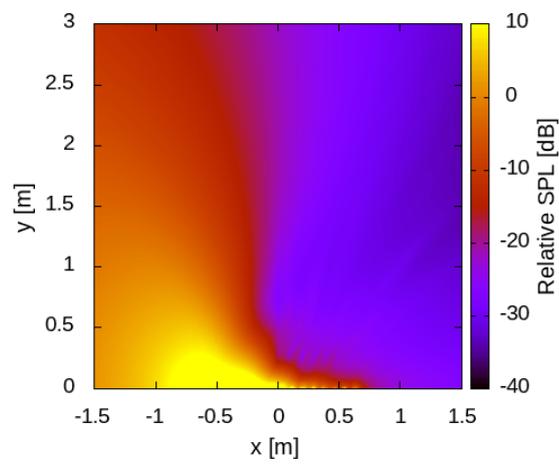
図 5.6 と図 5.7 に PESQ マップの結果を示す。PESQ はスピーカへの入力信号を原音とし，対象となる位置で合成された再生音声を評価した。図から，すべての配置において，再生エリアで PESQ 値が 4 程度の高い値を示しており，再生品質に問題がないことが確認できる。



(a) 格子配置

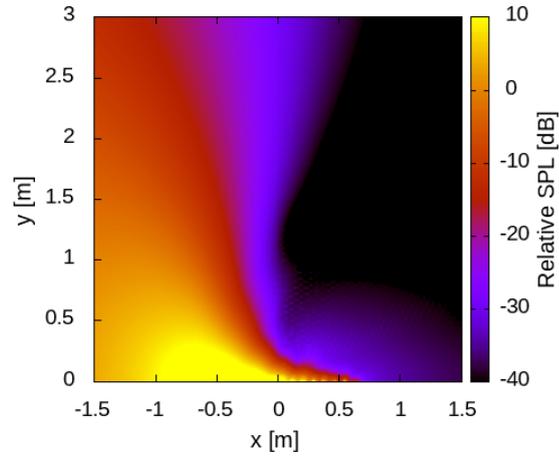


(b) 境界配置

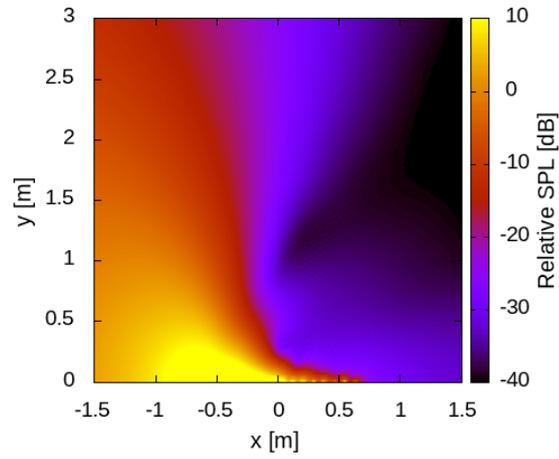


(c) L型配置

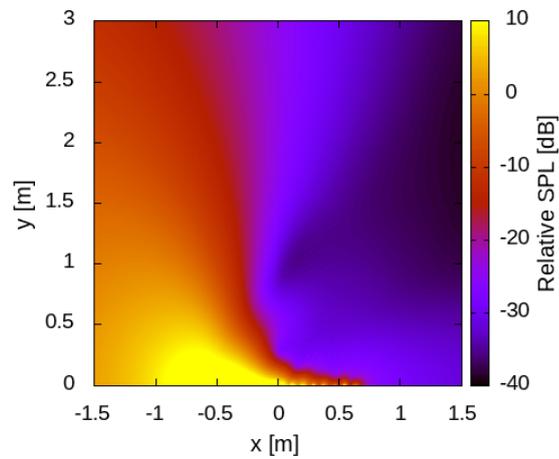
図 5.4: 各配置の相対音圧レベル (抑圧制御点間隔 0.1 m)



(a) 格子配置

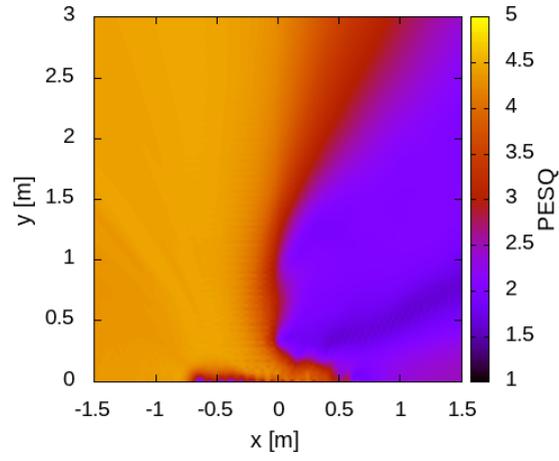


(b) 境界配置

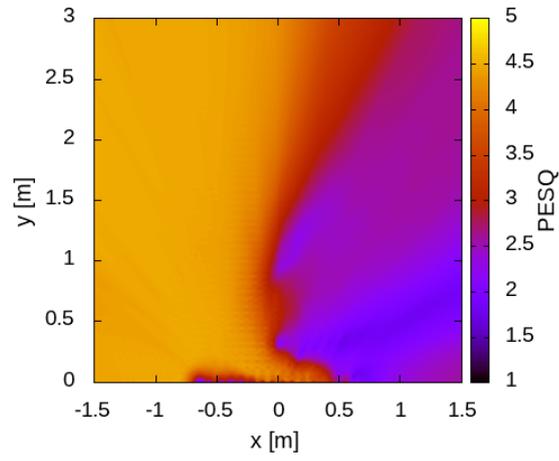


(c) L型配置

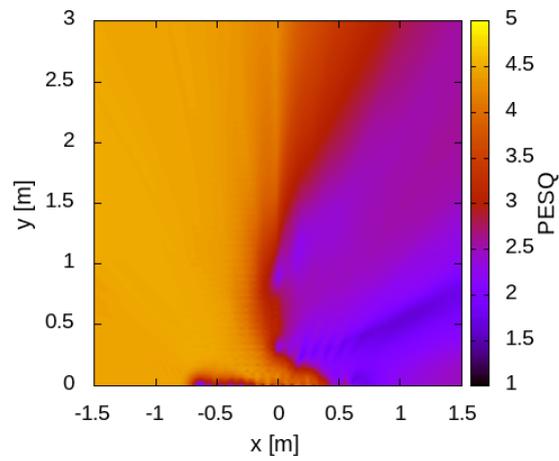
図 5.5: 各配置の相対音圧レベル (抑圧制御点間隔 0.02 m)



(a) 格子配置

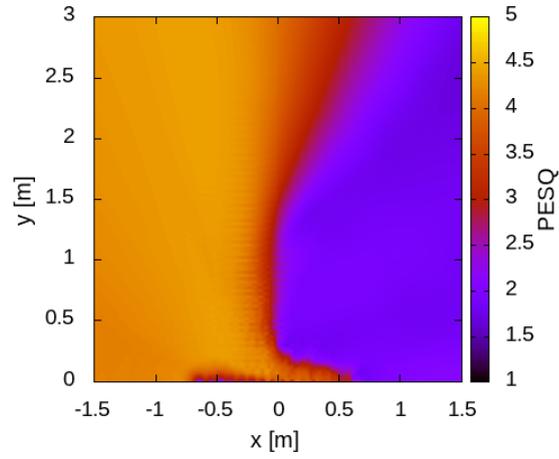


(b) 境界配置

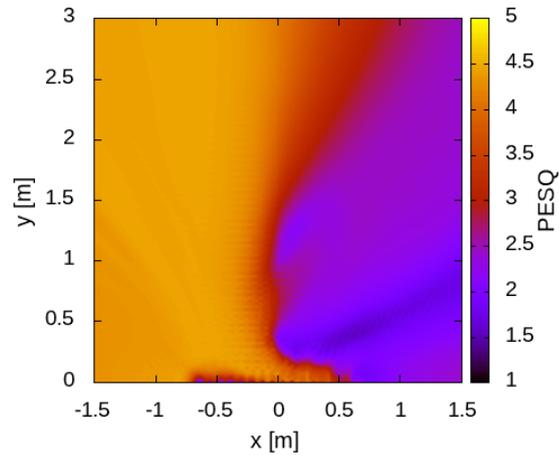


(c) L型配置

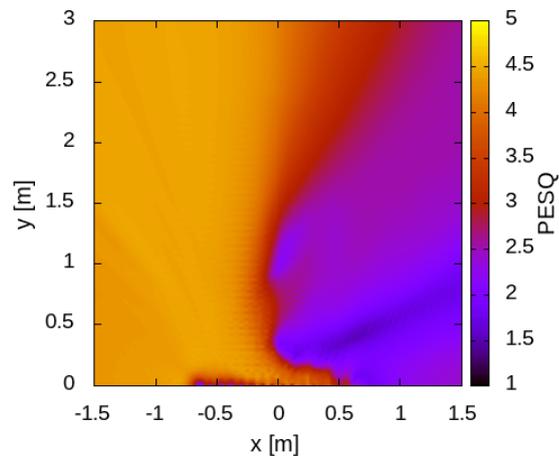
図 5.6: 各配置の PESQ (抑圧制御点間隔 0.1 m)



(a) 格子配置



(b) 境界配置



(c) L型配置

図 5.7: 各配置の PESQ (抑圧制御点間隔 0.02 m)

## 5.5 実環境実験

計算機シミュレーションの結果から、抑圧制御点が比較的少なく、良好な性能を示した 0.02 m 間隔の L 型配置を用いて、防音室における実環境実験を行った。実験はスピーカアレーの正面に設置した直線マイクロフォンによる音圧の測定と、単語了解度試験による実験を行った。

### 5.5.1 実験条件

図 5.8 に示す龍谷大学瀬田キャンパス 7 号館内にある残響時間約 140 msec の防音室において実験を行った。図 5.9 に作成した 16ch スピーカアレーを示す。使用したスピーカは Fostex 社の PM0.1 であり、スピーカ間隔は 0.085 m である。各スピーカから各制御点までのインパルス応答は  $2^{20}$  サンプルの Log-SS 信号 [61] 再生し、JTS 社の CX-500 を用いて収録した。インパルス応答長は 2048 サンプルとした。

### 5.5.2 伝達関数の違いによる音圧の変化

使用する伝達関数の違いによる性能の変化を調べるため、多点制御フィルタ算出の際に用いた伝達関数として、式 (2.13) の自由空間を仮定した伝達関数 (Free)、すべての制御点でインパルス応答を測定したもの (All) を比較する。また、スピーカアレーの両端のスピーカからそれぞれ異なる音声を再生するもの (1ch) についても実験を行う。

図 5.10 に Area A に再生を行ったときの相対音圧レベルを、図 5.11 に Area B に再生を行ったときの相対音圧レベルを示す。横軸は  $x$  方向の座標であり、スピーカ前方  $y = 1.0$  m の結果である。したがって、図 5.10 は  $x = -1.0$  m、図 5.11 は  $x = 1.0$  m が再生制御点となり基準点の 0 dB となっている。図 5.10 と図 5.11 から、再生エリアに対して抑圧エリアは約 15 dB 音圧が小さくなっており、それぞれエリア再生が実現できていることがわかる。また、再生制御点のインパルス応答を測定したもの (All) は自由空間伝達関数 (Free) と比べ、抑圧エリアの境界付近 ( $x = 0$  dB) で音圧が低くなっていることがわかる。しかし、その差は最大でも約 4.2 dB 以内である。

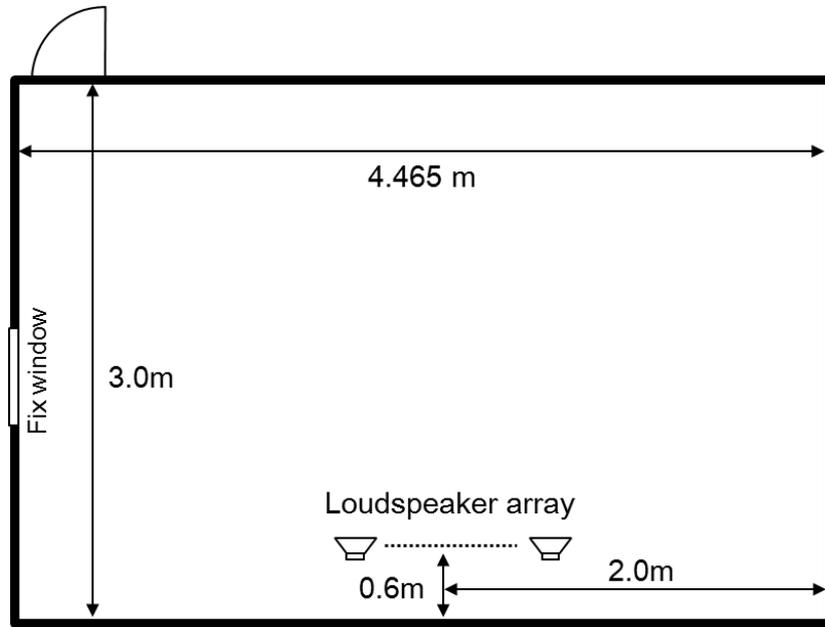


図 5.8: 実験を行った防音室の見取り図



図 5.9: 作成した 16ch 直線スピーカアレー

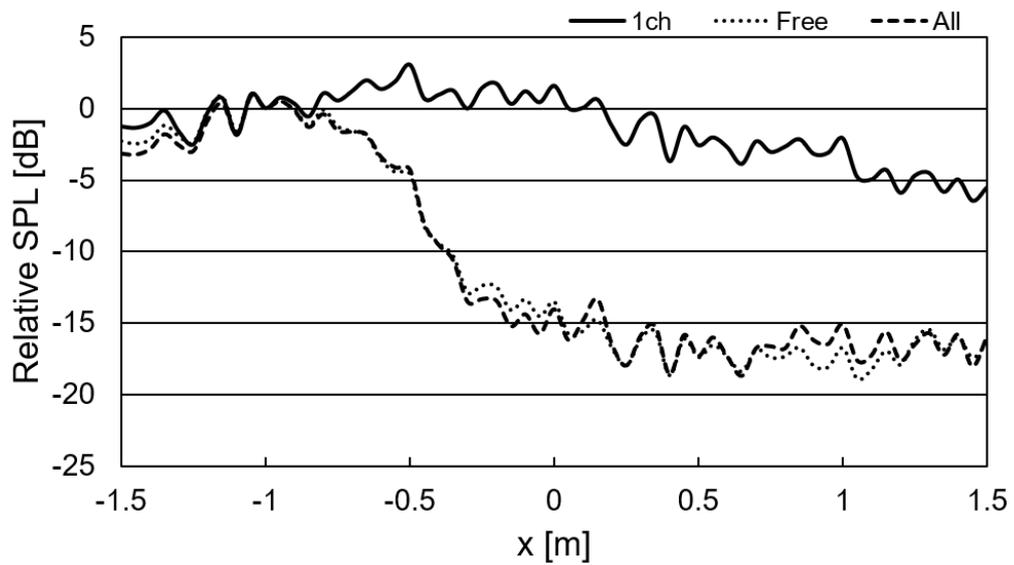


図 5.10: Area A に再生を行った時の相対音圧レベル

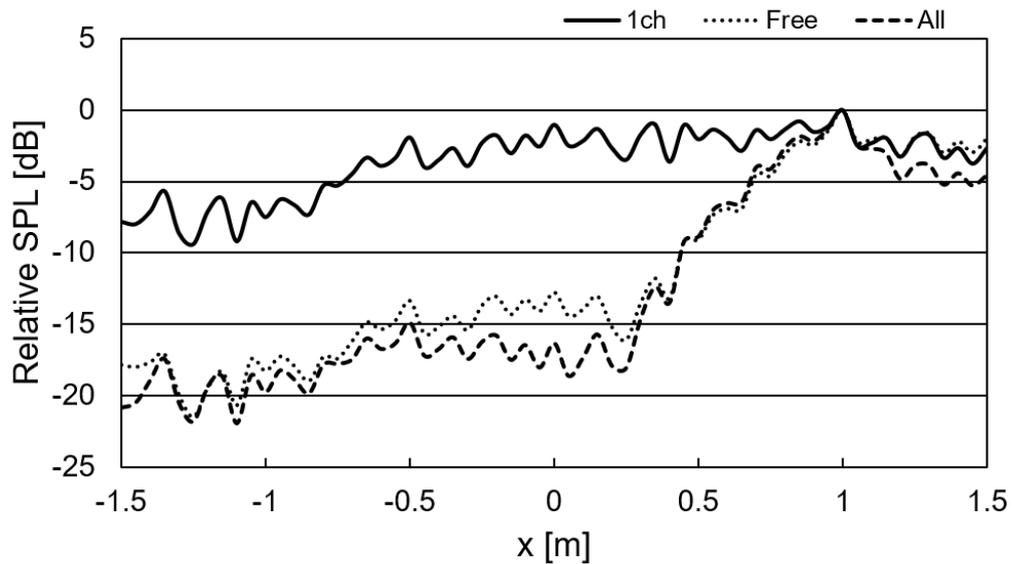


図 5.11: Area B に再生を行った時の相対音圧レベル

### 5.5.3 単語了解度試験

再生エリアでの再生品質と抑圧エリアでの音質を評価するため単語了解度試験による主観評価実験を行った。単語了解度試験は4 moraの単語を受聴し、書き取りを行う試験である。書き取った単語の正答率により複数エリア再生の評価を行う。

図5.12に試験時の配置を示す。聞き取る単語はArea Aの再生制御点に再生を行う。Area Bの再生制御点にはArea Aの単語とは異なる話者の8 secの発話音声を再生し、2つのエリアに異なる音声が再生されている状態とする。被験者はArea AとArea Bの再生制御点上(Listening position A, Listening position B)のいずれかに着席し、Area Aに再生された単語を聞き取って回答用紙に記入する。回答時間は約7秒とした。Area Aでは再生された単語が正しく聞き取れるかどうかを評価し、Area BではArea Aから漏えいする音声の単語が聞き取れないかどうかを評価する。

#### 試験条件

これまでの結果から、試験に用いる手法はL型配置で抑圧制御点間隔は0.02 mのものとした。多点制御法に使用するインパルス応答は、すべての制御点で自由空間のもの(Free)、すべての制御点で測定したもの(All)、アレーの両端のスピーカからそれぞれの音声を再生したもの(1ch)の3種類とした。試験には親密度別単語了解度試験用音声データセット(FW03)[47]の中から、親密度が最も高い5.5~7.0のものを使用した。評価数は各手法に対して、男女各2名の異なる4単語ずつとし、1条件当たり16単語を評価した。また、Area Bへ再生する音声はArea Aに再生する話者とは異なる男女各2名の16音声とした。Area Aに再生される単語とArea Bに再生される文章の男女ペアは総当たりとし、手法間で条件が偏らないように調整を行った。

試験は同じ防音室で行った。被験者は20代の男女計18名である。提示音圧は図5.12の受聴位置Aにおいて $L_{Aeq} = 50$  dBとした。再生を行う周波数帯域は300 Hz~3.4 kHzである。また、受聴位置Bの被験者がArea Aに再生した単語ではなくArea Bに再生した音声を書き取らないように試験前に説明を行い、練習を十分実施した。しかし、そ

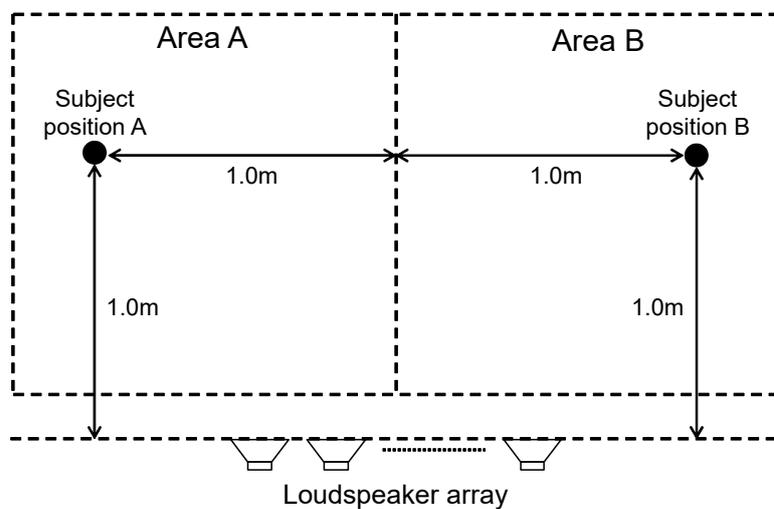


図 5.12: 単語了解度試験時の受聴者の位置

れであっても Area B に再生した音声を記入した場合は聞き取れていないとみなし，不正解とした。

### 試験結果

図 5.13 に受聴位置 A での結果，図 5.14 に受聴位置 B での結果を示す。図中のエラーバーは 95%信頼区間を示す。図 5.13 から伝達関数が Free, All とともに約 90%の正答率となっており，聞き取りに問題のないことが確認できる。一方，図 5.14 から，受聴位置 B では 1ch 再生と比べて，多点制御法用いた場合の正答率が低くなっている。この結果から，多点制御法によって，Area A からの音声の漏えいが抑えられていることがわかる。さらに，その中でも All が約 9%の正答率となっていることから，実際のインパルス応答を測定することで，より抑圧できており，L 型の制御点配置であっても十分な抑圧量を確保できると考えられる。また，自由空間伝達関数を用いた Free の場合の正答率が 40%程度であることから，使用する状況によっては十分実用的であると考えられる。

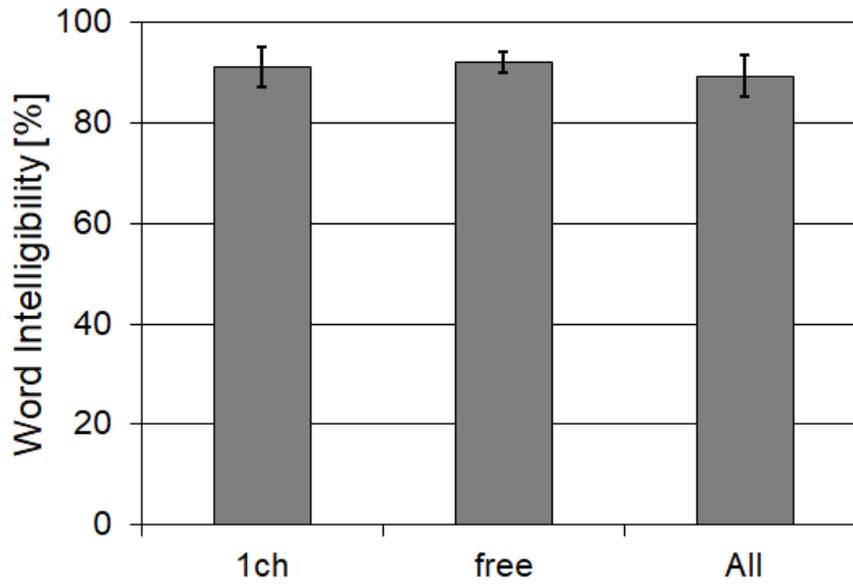


図 5.13: 受聴位置 A での単語了解度試験結果

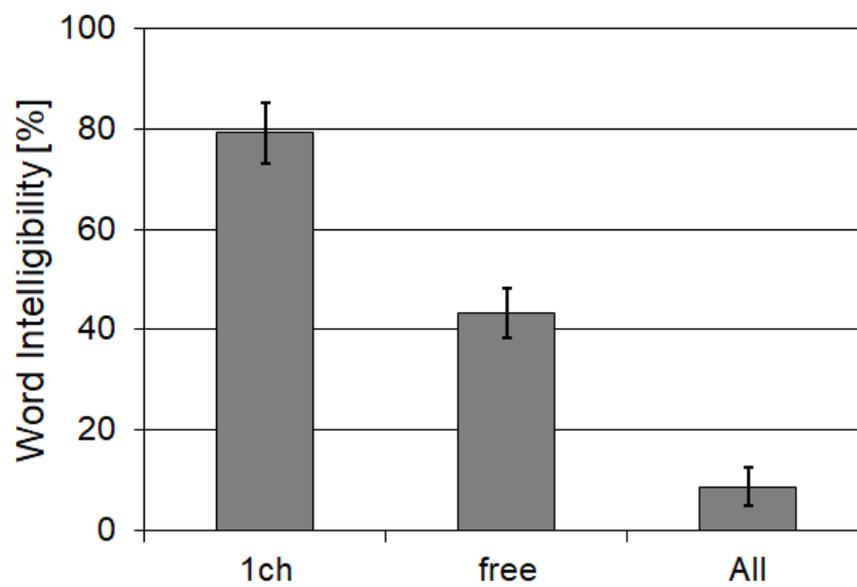


図 5.14: 受聴位置 B での単語了解度試験結果

## 5.6 まとめ

本章では，1つのスピーカアレーを用いて複数のエリアに異なる音声を再生する複数エリア再生について検討を行った．少ないスピーカ数，制御点数で広い範囲を制御するためには，制御点の配置が重要である．本章では，抑圧エリアを形成するための抑圧制御点配置として，L字型の制御点配置を提案した．

計算機シミュレーションから，L型の制御点配置を用いた場合であっても領域境界を囲う配置と同等の抑圧性能を示した．また，PESQによる音質評価を行い，すべての配置で再生エリアでの音質に問題がないことを確認した．これらの結果から，L型配置の抑圧制御点間隔 0.02 m を用いて実環境での評価実験を行った．実環境実験では，多点制御法に用いる伝達関数について実測値と自由空間伝達関数を用いる場合の比較を行った．結果から，再生エリアと比べて抑圧エリアは約 15 dB の抑圧を実現できた．さらに，単語了解度試験において，実測した伝達関数を用いた場合の正解率が約 9 %，自由空間伝達関数を用いた場合の正解率が約 40 % となり，使用する環境によっては自由空間伝達関数を用いた場合であっても，十分実用的であることを示した．

# 第6章 遺伝的アルゴリズムを用いた 制御点配置の決定法

## 6.1 はじめに

多点制御法によるエリア再生は、各制御点での音圧特性を制御するため、制御点の位置を決定する必要がある。本研究では再生制御点は固定し、抑圧制御点の位置を遺伝的アルゴリズムによって決定する。特に前章で述べたような、ある程度広い範囲に再生を行う場合は、部屋の大きさやスピーカアレーの形状などに応じてその都度、抑圧制御点の配置を決定する必要がある。

本章では制御点配置の決定を組合せ最適化問題とみなし、遺伝的アルゴリズムによる最適化手法の適用を検討する。以下、6.2章において、遺伝的アルゴリズムについて説明する。6.3節において、多点制御法に遺伝的アルゴリズムを適用するための遺伝子表現について説明する。6.4節において、制御点を決定するための遺伝的アルゴリズムの詳細と評価関数について述べる。6.5節において、計算機シミュレーションによる評価実験の概要と結果について述べる。6.6節において、本章のまとめを述べる。

## 6.2 遺伝的アルゴリズム

遺伝的アルゴリズム (GA: Genetic Algorithm) は特定の問題に限らず幅広い組合せ最適化問題を解く手法として知られている [62, 63]。生物の進化過程を模擬した遺伝的アルゴリズムは、進化的計算手法の中で最も一般的な計算手法である。データを遺伝子で表現した個体を複数個用意し、選択、交叉、突然変異といった遺伝的操作を繰り返すことで解を求める。

## 6.3 制御点の遺伝子表現

本研究では、各制御点の x-y 座標を遺伝子  $(x_i, y_i)$  とする。遺伝的アルゴリズムは変化させる変数を遺伝子とし、その遺伝子の集合を染色体として表現する。したがって、1 個体は 1 つの制御点配置を表し、その染色体は式 (6.1) のようになる。

$$\mathbf{V} = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \\ y_1 & y_2 & \cdots & y_N \end{bmatrix} \quad (6.1)$$

ここで、 $(x_1, y_1)$  は再生制御点の位置、 $N$  は制御点数である。再生制御点を表す遺伝子  $(x_1, y_1)$  は進化によって変化しない。

## 6.4 遺伝的アルゴリズムによる制御点配置の最適化

図 6.1 に遺伝的アルゴリズムの流れを示す。初めに、抑圧制御点をランダムに配置した個体を初期集団として複数個生成する。この集団を親集団とする。次に、各個体の適合度を計算し、その結果に基づいて 2 個体を選択する。適合度は 6.4.3 節で述べる評価関数を用いて計算する。選択された個体は交叉確率と突然変異確率に基づき交叉、突然変異を行う。この操作を親集団と同じ数になるだけ繰り返し子集団とする。子集団は次の世代の親集団となり、終了条件を満たすまで繰り返す。適合度の計算から子集団の生成までを 1 世代とする。

### 6.4.1 個体の選択方法

親集団から 2 個体を選択する方法として、ルーレット選択 (適合度比例選択)[64] を用いる。ルーレット選択は各個体の適合度に基づき、ルーレットを作成し、選択を行う手法である。適合度が高い個体ほど、ルーレットの面積が大きくなり、選ばれる確率が高くなる。しかし、適合度が低い個体であっても、選択確率はゼロにならないため、比較的局所解に陥りにくい手法である。

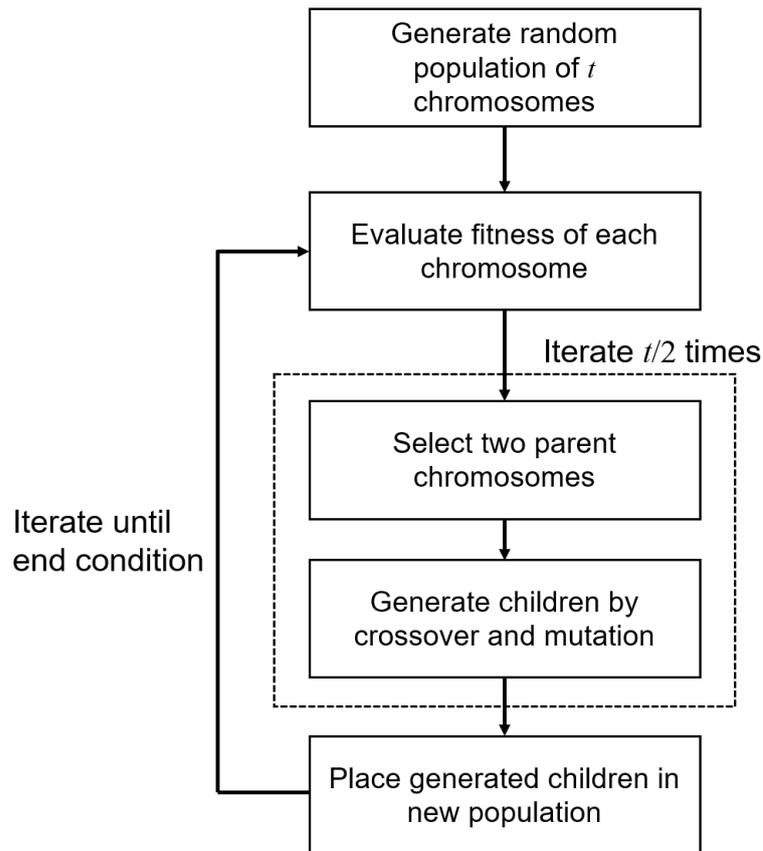


図 6.1: 遺伝的アルゴリズムの処理の流れ

ルーレット選択は解が収束に近づくにつれて各個体の適合度が近くなり、良い個体を選ばれ難くなる。そこで、線形スケールリング [64] を導入し適合度に差をつける。本研究では線形スケールリングの係数は2とした。

#### 6.4.2 交叉と突然変異

交叉は一様交叉を用いる。図 6.2 に染色体長  $N=3$  の時の例を示す。一様交叉は染色体長のビットマスクをランダムに作成し、子1の場合、ビットの値が0ならば親1の遺伝子を受け継ぎ、ビットが1ならば親2の遺伝子を受け継ぐ。子2は、ビットの値が0ならば親2の遺伝子を受け継ぎ、ビットが1ならば親1の遺伝子を受け継ぐ。交叉するかどうかは交叉確率に基づく。交叉しないと判定された場合、親の2個体そのまま子の2個体となる。

$$\begin{array}{l}
 N = 3 \quad \text{Parent1} \qquad \qquad \text{Parent2} \\
 V_1 = \begin{pmatrix} 3 & 6 & 4 \\ 1 & 3 & 2 \end{pmatrix} \qquad V_2 = \begin{pmatrix} 2 & 3 & 5 \\ 4 & 0 & 3 \end{pmatrix} \\
 \text{Masking bit array} \quad \boxed{0} \quad \boxed{1} \quad \boxed{1} \\
 \text{If bit is 0, select from Parent1} \quad \text{If bit is 0, select from Parent2} \\
 \text{If bit is 1, select from Parent2} \quad \text{If bit is 1, select from Parent1} \\
 \\
 \text{Child1} \qquad \qquad \qquad \text{Child2} \\
 v_1 = \begin{pmatrix} 3 & 3 & 5 \\ 1 & 0 & 3 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 2 & 6 & 4 \\ 4 & 3 & 2 \end{pmatrix}
 \end{array}$$

図 6.2:  $n = 3$  としたときの一様交叉の例

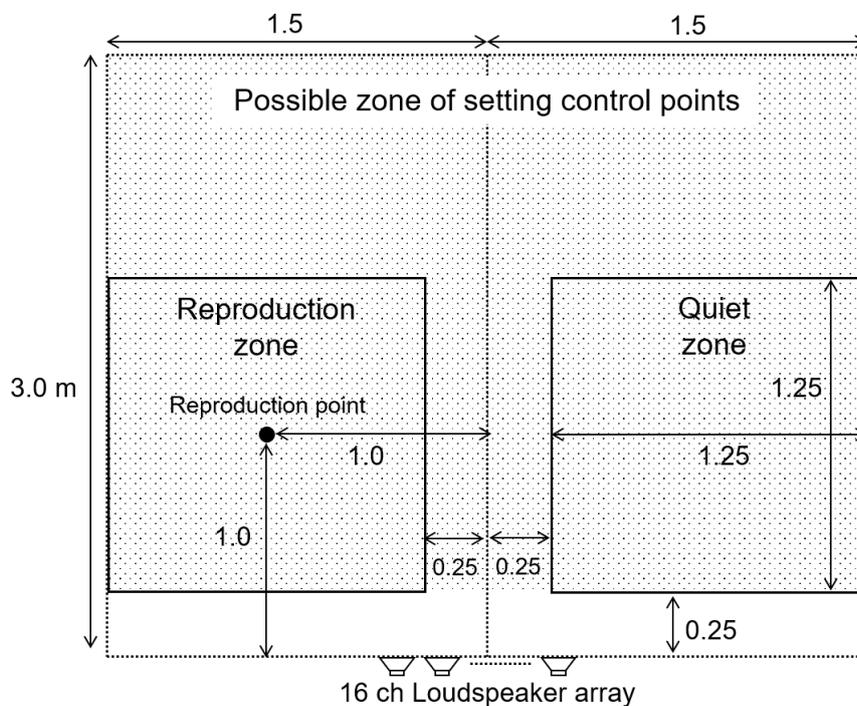


図 6.3: 各エリアとスピーカアレーの配置

突然変異は、子の1つの遺伝子についてランダムな値と置き換える。突然変異を行うかどうかは突然変異確率に基づき、それぞれの子について判定を行う。交叉と突然変異は独立しており、交叉したかどうかは突然変異に影響を及ぼさない。

### 6.4.3 評価関数

図 6.3 に本研究における各エリアと抑圧制御点を取りうる範囲を示す。制御範囲はスピーカアレーの前方 3 m × 3 m とし、再生エリアは左側の 1.25 m × 1.25 m、抑圧エリアは右側の 1.25 m × 1.25 m とする。抑圧制御点を配置できる範囲は図 6.3 網掛け部分の 3 m × 2.75 m とした。

本研究では、適合度  $F$  は式 (6.2) で表されるエリア間の音圧比評価関数を用いる。

$$F = 10 \log_{10} \frac{\sum_{\epsilon_j} |P_j|}{\sum_{\epsilon_k} |P_k|} \quad (6.2)$$

$P$  は各評価点での音圧、 $j$  は再生エリア内の評価点、 $k$  は抑圧エリア内の評価点である。したがって、再生エリアと抑圧エリアの音圧比が大きくなるほど、エリア再生が実現できており、良い評価となる。

式 (6.2) は音圧比のみを評価しているため、再生品質について考慮していない。多点制御法は、制御点での音圧特性を周波数領域で制御するため、制御点配置が音質に影響を及ぼすことが考えられる。そこで、式 (6.2) の評価関数に PESQ を導入する。PESQ は MOS (Mean Opinion Score) 値を客観的に推定するアルゴリズムであり、MOS 値のカテゴリ (5:非常に良い, 4:良い, 3:普通, 2:悪い, 1:非常に悪い) に対応する。PESQ を用いることで音質を考慮した適合度の計算を行うことができる。PESQ を含めた評価関数を式 (6.3) に示す。

$$F' = F + A \cdot \frac{1}{L} \sum_{l=1}^L PESQ(S_l) \quad (6.3)$$

ここで、 $A$  は加算割合を調整する重み係数、 $S_l$  は計測点での信号、 $L$  は計測点の個数である。再生エリア内での PESQ 計測点の平均値を従来の評価関数に加算することで適合度とする。

## 6.5 計算機シミュレーション

突然変異確率を決定するための予備実験と，提案手法の性能を評価するため計算機シミュレーションを行った．本節では実験条件とそれぞれの結果について述べる．

### 6.5.1 実験条件

多点制御法の制御点数は再生制御点 1 点，抑圧制御点 31 点とし，式 (2.8) の正則化パラメータは周波数によらず 0.1 とした．スピーカ数は 16，スピーカ間隔は 0.085 m とした．式 (2.5) の伝達関数については式 (2.13) の自由空間伝達関数を用いる．

遺伝的アルゴリズムの初期個体数は 32 とし，交叉確率は 0.6 とした．終了条件は適合度の最大値が 15 世代連続して性能の向上がない場合，または適合度の平均値と中央値が一致した場合とした．

式 (6.2) の音圧比については 100~7000Hz のホワイトノイズ，式 (6.3) の PESQ 値については内容の異なる 4 秒間の発話音声を男女各 4 名，計 32 秒を用いて計算を行った．PESQ は ITU-T P.862.2 で定義される広帯域 PESQ とし，リファレンスは入力信号とした．音圧比と PESQ の測定点を図 6.4 に示す．音圧比はそれぞれのエリア全体において 0.25 m 間隔で各 36 点，PESQ は再生エリアのみに計測点を配置し，再生制御点上とその周囲 8 点の合計 9 点とした．

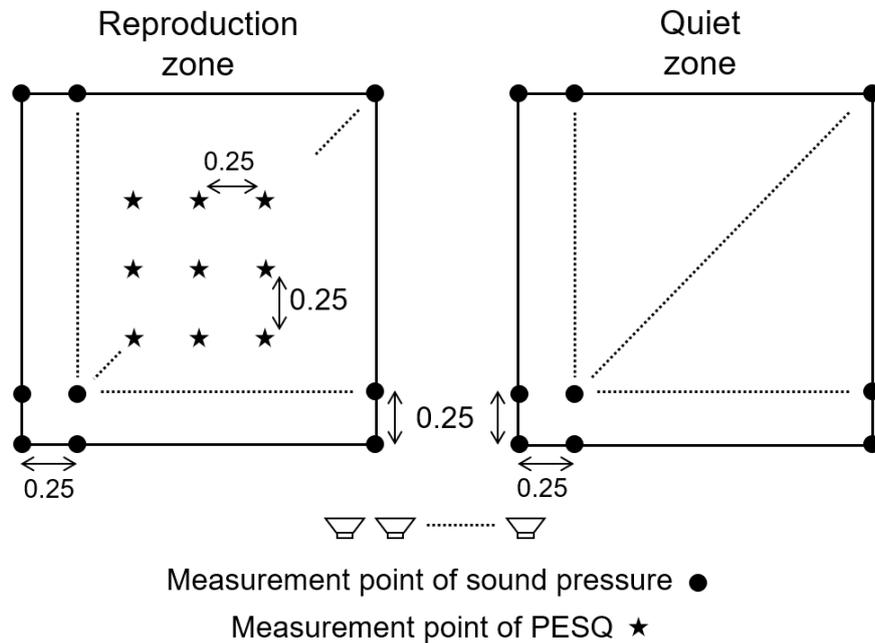
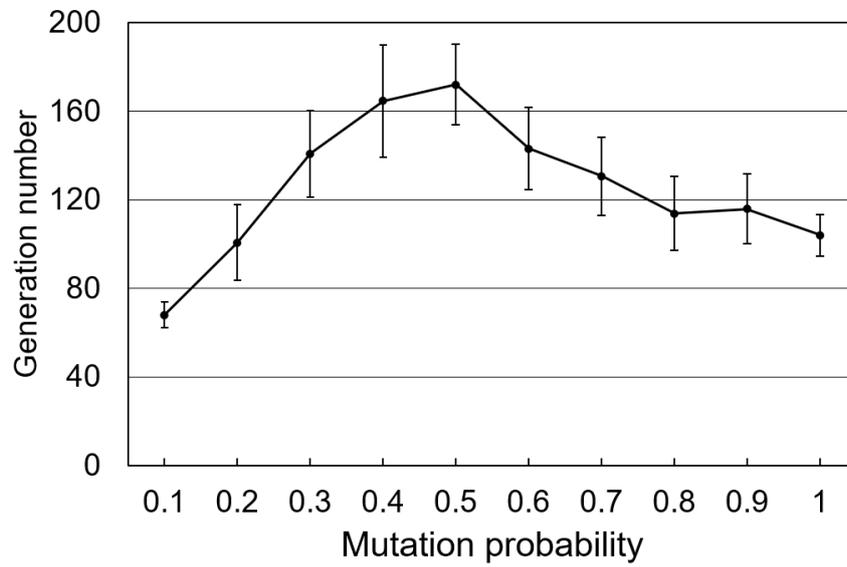


図 6.4: 音圧と PESQ の測定点

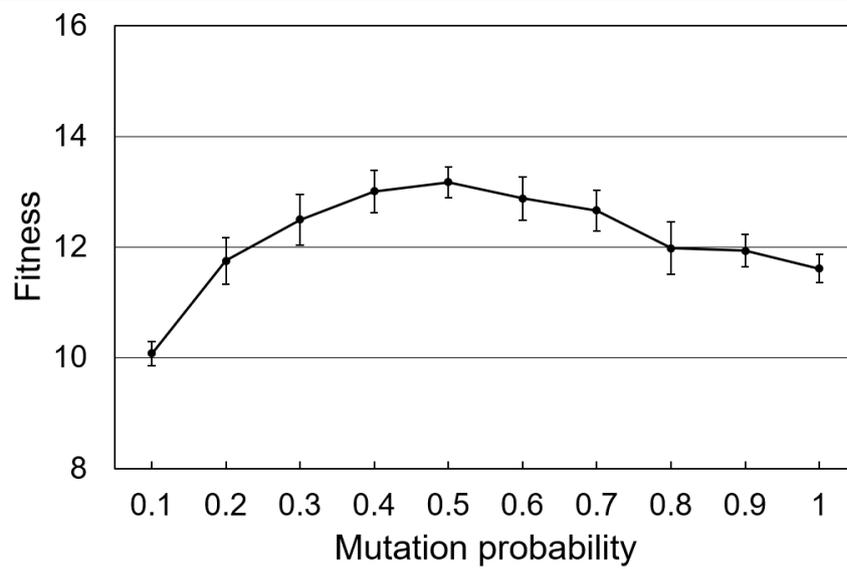
### 6.5.2 突然変異確率の検討

通常の遺伝的アルゴリズムでは突然変異確率を 0 に近い小さい値に設定する。しかし、本研究では制御点の座標を遺伝子としており、一様交叉はランダムに選ばれた制御点をそれぞれの親同士で交換することに相当する。抑圧制御点は制御点同士の位置関係が重要であるため、初期配置によっては局所解に陥る可能性がある。そこで、突然変異によって多様な解を生成し、局所解から抜け出すことを考える。図 6.5 に突然変異確率を変化させたときの収束に要する平均世代数と適合度を示す。PESQ 係数は  $A = 0$  とし、それぞれ 20 回試行した平均値である。図中のエラーバーは 95% 信頼区間を示す。

図 6.5 から突然変異確率が 0.5 の時に収束に要する世代数が最も多く、また、適合度が最も高くなっていることが確認できる。これらの結果から、本研究では突然変異確率は 0.5 とする。



(a) 世代数



(b) 適合度

図 6.5: 突然変異確率を変化させたときの収束世代数と適合度

### 6.5.3 実験結果

図 6.6 に PESQ の重み係数  $A$  に対する音圧比と PESQ の変化を示す。値はそれぞれの重み係数において 15 回試行した平均値であり、エラーバーは 95%信頼区間を表す。図から重み係数を 5 以上に設定することで、最終的な PESQ の値が上昇しており、再生エリアでの音質が向上したと考えられる。また、図 6.7~図 6.9 に重み係数を 0, 5, 10 としたときの、最終的な適合度が最も良い試行での世代ごとの音圧比と PESQ の変化を示す。図から、 $A=10$  では世代が進むにつれて PESQ が上昇していることがわかる。しかし、重み係数が大きくなると音圧比は低下し、エリア再生の精度が低下していることから、エリア再生の性能と再生制御点での音質はトレードオフになっていることがわかる。

図 6.10-図 6.12 に図 6.7-図 6.9 の試行での 1 世代目、100 世代目、終了時の制御点配置と相対音圧マップを示す。赤丸が再生制御点、青丸が抑圧制御点である。図より、どの係数であっても終了時には抑圧エリア側に抑圧制御点が集まっている。また、重み係数  $A=0$  は抑圧エリア内のスピーカアレーに近い位置に密集しているのに対し、 $A=10$  では抑圧エリア内にまばらに位置しており、一部の制御点は再生エリアを囲むように配置されている。これは再生制御点周辺に強い指向性を形成することで、音質の向上が期待できるためであると考えられる。

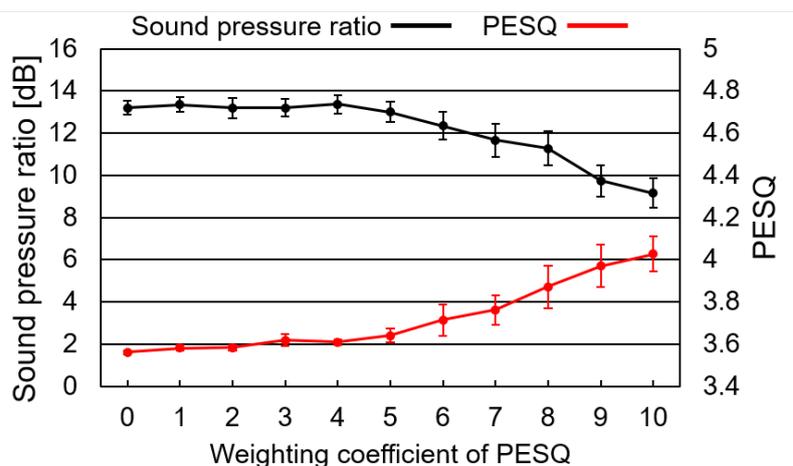


図 6.6: PESQ の重み係数による相対音圧レベルと PESQ の平均

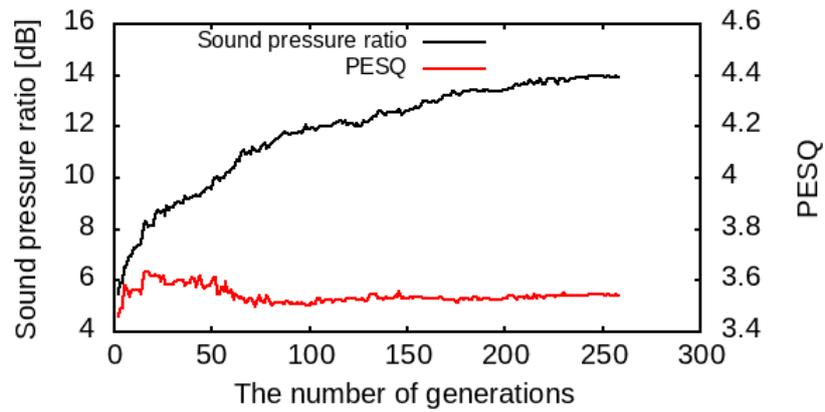


図 6.7: 各世代ごとの相対音圧レベルと PESQ の変化 ( $A = 0$ )

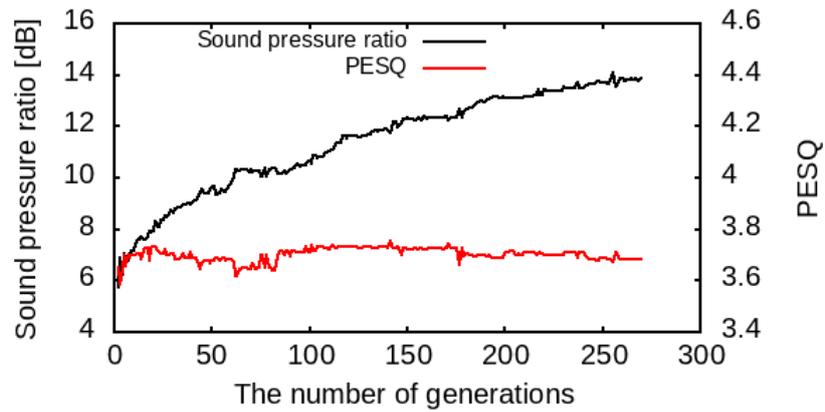


図 6.8: 各世代ごとの相対音圧レベルと PESQ の変化 ( $A = 5$ )

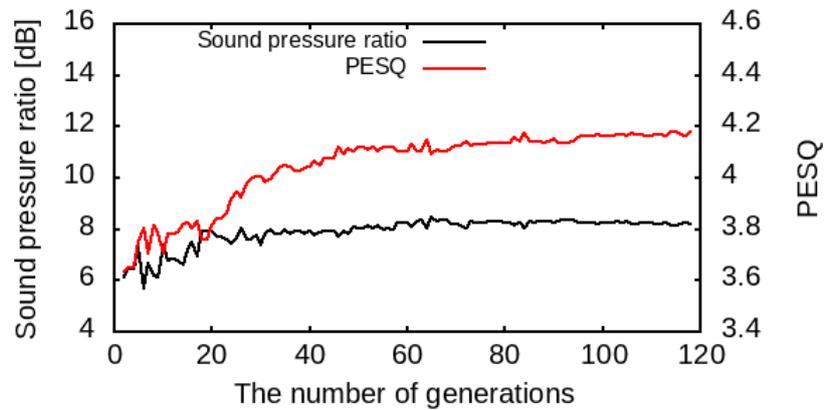
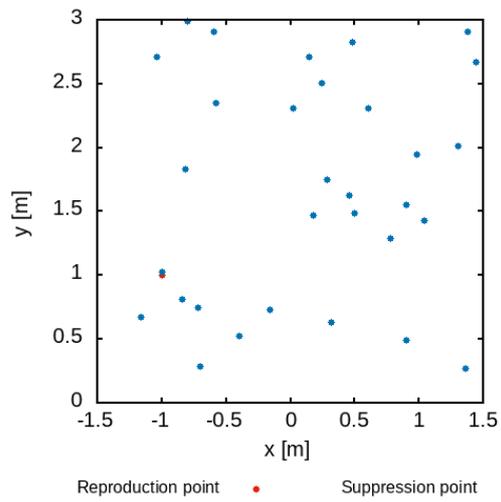
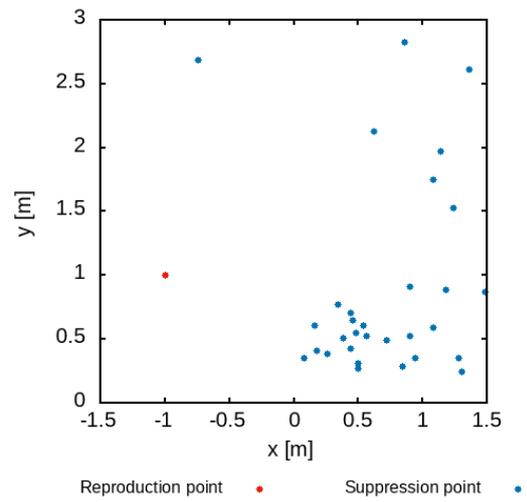


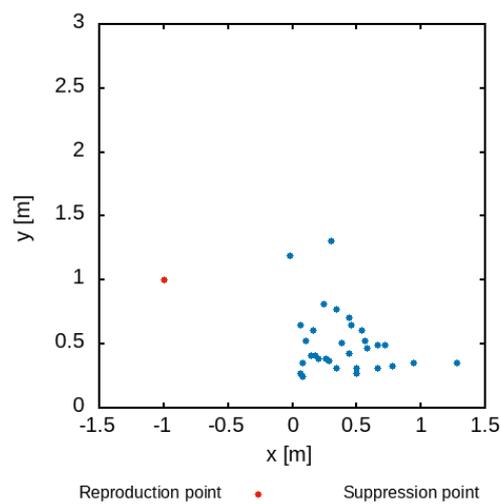
図 6.9: 各世代ごとの相対音圧レベルと PESQ の変化 ( $A = 10$ )



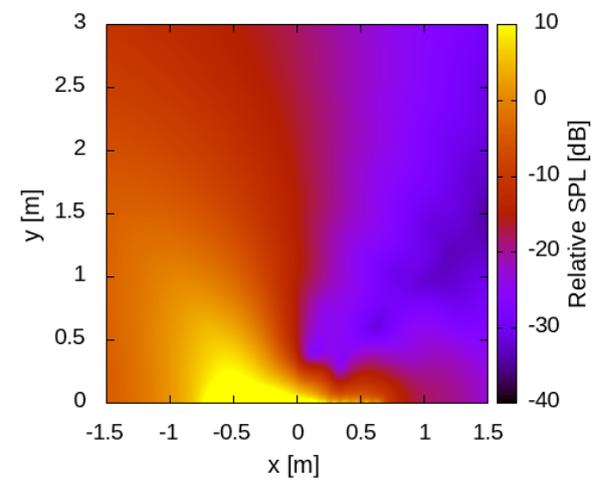
(a) 1 世代目



(b) 100 世代目

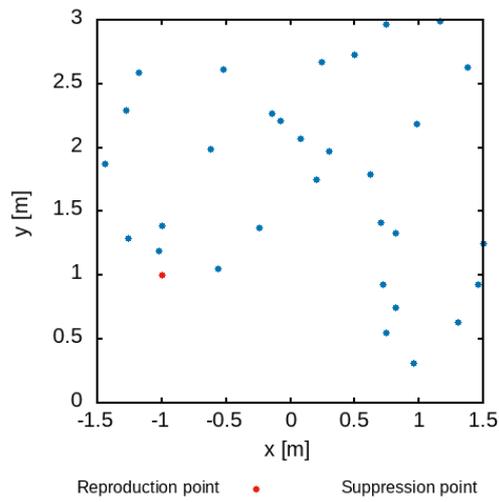


(c) 最終世代 (258 世代目)

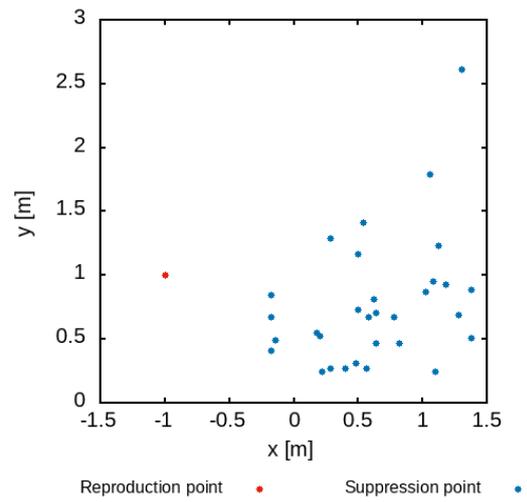


(d) 最終世代 (258 世代目) の相対音圧マップ

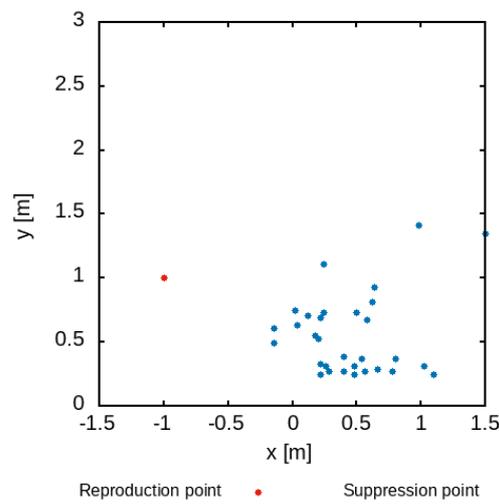
図 6.10: 各世代の制御点配置と最終世代の相対音圧マップ ( $A = 0$ )



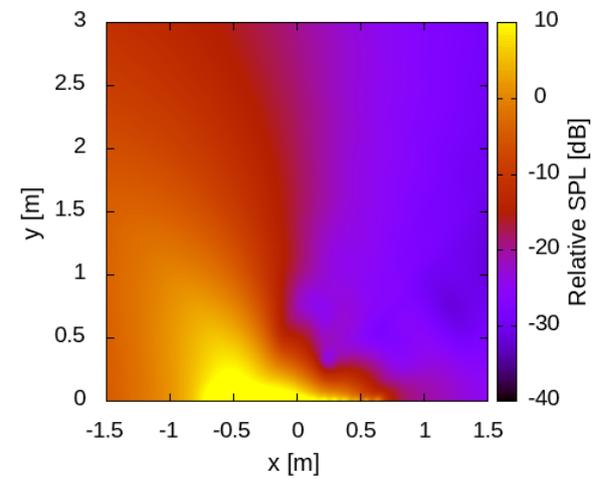
(a) 1 世代目



(b) 100 世代目

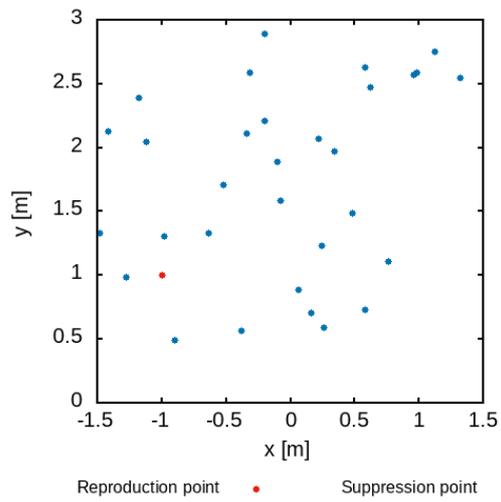


(c) 最終世代 (270 世代目)

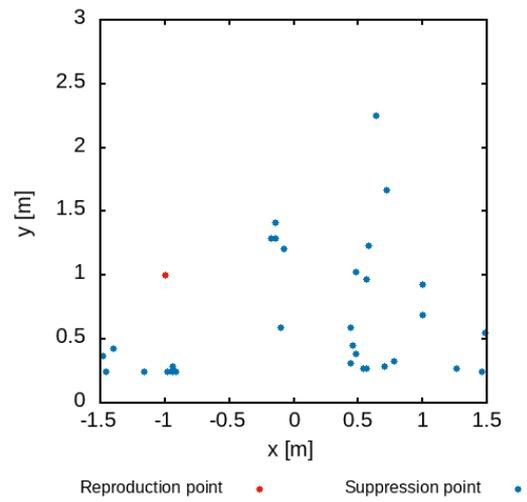


(d) 最終世代 (270 世代目) の相対音圧マップ

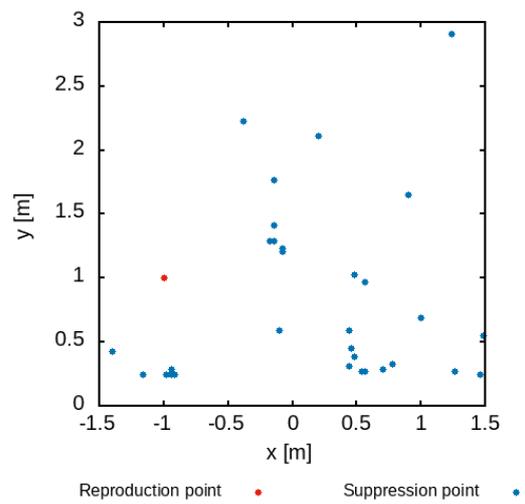
図 6.11: 各世代の制御点配置と最終世代の相対音圧マップ ( $A = 5$ )



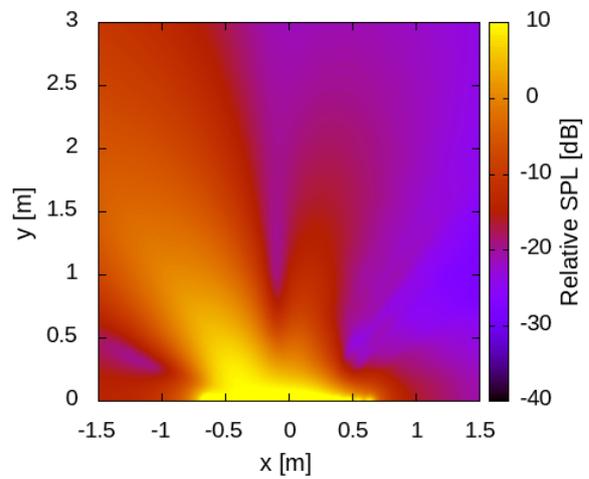
(a) 1 世代目



(b) 100 世代目



(c) 最終世代 (118 世代目)



(d) 最終世代 (118 世代目) の相対音圧マップ

図 6.12: 各世代の制御点配置と最終世代の相対音圧マップ ( $A = 10$ )

## 6.6 おわりに

本章では、遺伝的アルゴリズムを用いて複数エリア再生のための制御点配置を決定する手法を提案した。また、評価関数に再生制御点周辺の PESQ 値を導入することで、音質を考慮した制御点配置の決定法について提案した。計算機シミュレーションによる評価実験から、得られた抑圧制御点は抑圧エリアのスピーカアレーに近い境界に集まる傾向があることを確認した。また、PESQ 係数を大きくし、より音質を重視する設定にした場合は、抑圧制御点が再生エリア周辺にも分布し、再生制御点に向けて指向性を形成するような配置になることを確認した。

今回提案した遺伝的アルゴリズムによる制御点配置の決定手法では、制御点の座標をそのまま遺伝子として扱い、一様交叉を行っている。しかし、多点制御法は制御点同士の位置関係が重要であり、一様交叉を行うことでその関係性が崩れ、結果として収束に時間がかかることがある。今後の課題として、制御点間の位置関係を考慮した最適化手法を検討する必要がある。

# 第7章 結論

## 7.1 本論文のまとめ

本論文は、多点制御法を用いた音のエリア再生に関する手法について、研究成果をまとめたものである。本研究では領域内に設定した制御点での音圧を直接制御する多点制御法を用いて、空間内の狭い範囲に再生する手法と、ある程度広範囲に再生し、複数のエリアを形成する手法について検討を行った。空間内の狭い範囲に再生する手法では、信号分解法によって分解したサブ信号を多点制御法によって空間上で再合成する手法を検討した。また、目的のエリア外に漏えいした音声に対してマスキングによるスピーチプライバシー保護手法を導入し、音声内容の聞き取り難さを上昇させる検討を行った。複数エリア再生については、ブロードサイド型のスピーカアレーを用いて、抑圧エリアでの制御点配置の検討を行い、遺伝的アルゴリズムによる配置の最適化について検討を行った。

第2章では、複数のスピーカを用いて音場を制御する手法とその原理を応用したエリア再生の手法について述べた。音場再現の手法として、領域境界の音圧を制御する境界音場制御と領域内の音圧を一致させる多点制御法について説明した。また、多点制御法を用いて、空間内に音圧が大きい場所と小さい場所を生成するエリア再生の手法を説明した。相対音圧レベルによる評価と主観評価実験を行い、再生を行うエリア外において音声内容が聞き取れる可能性があることを示した。

第3章では、音声聞き取れる範囲を狭くするため、信号分解法と多点制御法を組み合わせたエリア再生手法を提案した。入力信号を聞き取れないサブ信号に分解し、それぞれを多点制御法によって同じ場所に再生を行うことで、その場所ではサブ信号が再合成されて聞き取れるようになり、それ以外の場所ではサブ信号の位相がずれるこ

とで、聞き取りを困難にすることができる。また、多点制御法を組み合わせることで、目的のエリア外で音圧を低下させ、不快感の低減を検討した。サブ信号の生成手法として、従来手法であるランダムベクトル分解法を説明し、目的方向での音質に問題がある可能性を示した。その解決方法として、入力信号をLPC分析によってLPC係数とLPC残差信号に分解し、残差信号を用いて複数のサブ信号を生成する手法を提案した。相対音圧レベルによる評価と主観評価実験から提案手法の有効性を確認した。

第4章では、多点制御法において指向性が十分に形成できず、周囲に漏えいする音声に対して、マスキングを行う手法を提案した。多点制御法を用いたエリア再生手法において、再生を行うエリア外で音声内容が聞き取れる問題があった。この問題に対して、再生対象外の領域にマスキング音を再生することで音声内容が聞き取れないようにするが、再生したマスキング音が再生対象のエリアに漏えいし、本来の再生エリアにおいて聞き取り難くなる可能性があった。そこで、マスキング音の再生についても多点制御法を用いて制御を行い、対象エリアへの漏えいを低減する手法を検討した。主観評価実験によって提案法の有効性を確認した。

第5章では、多点制御法を用いて、1つのスピーカアレーから異なる領域に異なる音声を再生する、複数エリア再生のための制御点配置について提案した。多点制御法は領域内に設置した制御点での音圧を制御することで指向性を形成する。そのため、制御点数が多くなると伝達関数の計測コストや逆行列演算による誤差が大きくなり、音声品質が低下する可能性があった。そこで、制御点を制御領域の境界とスピーカアレーに近い位置のみにL型に配置することを検討した。また、実際にスピーカアレーを作成し、実環境における実験を行った。実測した伝達関数を用いた場合と自由空間の伝達関数を用いた場合について相対音圧レベルによる比較した。さらに、単語了解度試験を行い、自由空間伝達関数を用いた場合であっても十分な性能が得られる可能性を示した。

第6章では、遺伝的アルゴリズムを用いて制御点配置を決定する手法を提案した。第5章で述べた複数エリア再生について、より効率的に制御を行うため、遺伝的アルゴリズムによる配置の決定方法について検討を行った。また、遺伝的アルゴリズムの評価

関数に、客観的な音質評価指標である PESQ を導入し、再生点付近での音質を考慮した制御点配置についても検討を行った。計算機シミュレーションによる評価実験を行い、エリア再生の性能の向上を確認した。また、得られた制御点配置が第 5 章で提案した制御点配置に近い配置であることを確認した。

## 7.2 今後の課題

本研究では多点制御法を用いたエリア再生について、エリア外に漏えいした音声の対策や制御点配置について検討を行ったが、さらに高精度で実用的な音のエリア再生を実現するためには以下の問題点が残されている。

1. 残響環境下での高精度なエリア再生
2. 領域内における人の移動への対応

1. に関しては、本研究で行った評価実験は、残響の無い自由空間を想定した計算機シミュレーションによる実験と残響時間が 140 msec と非常に短い防音室での実験である。したがって、一般室内のようなある程度の残響が存在する環境ではエリア再生の精度が低下すると考えられる。この問題を解決するためには、残響を含めた伝達関数を測定する必要があるが、人の移動など部屋の内部状況の変化が激しいことが予想できる。そのため、伝達関数を逐次的に測定するか、残響の影響を低減する手法が必要となる。例えば、スピーカアレーの近傍のみを制御し、急速に減衰させることで、反射の影響を無視できる手法が提案されている [65, 66]。このような手法を応用することで、一般の部屋でも精度の良いエリア再生が実現できると考えられる。

2. に関しては、本研究では多点制御法の再生制御点はあらかじめ固定してフィルタの設計を行っている。したがって、情報が欲しい人がその制御点位置まで正確に移動する必要がある。この問題を解決するためには、人の位置をセンサーやカメラを用いて認識し、その位置に応じた伝達関数を選択することが考えられる。人の追従ができれば、受聴位置の誤差を補正することや、歩行中の人を追従することが可能となり、デジタルサイネージなどにとって、より使いやすいシステムになると考えられる。

これらの問題を解決するために、今後も引き続き研究を行い、さらに高精度で使いやすいエリア再生手法の実現を目指す。

# 謝辞

本研究を行うにあたり、多くの方々にご指導とご鞭撻をいただきました。ここに感謝の意を表します。

龍谷大学理工学部 片岡章俊教授には、指導教員として筆者が研究室に配属された時から、研究の進め方をはじめ、研究内容、学会発表、論文執筆など丁寧にご指導いただきました。心から感謝いたします。また、本論文をまとめるにあたり、暖かいご指導、ご助言をいただいた龍谷大学理工学部 外村佳伸教授、藤田和弘教授に深く感謝いたします。

東京電機大学 金田豊教授、電気通信大学 羽田陽一教授には、スピーカアレー作成やインパルス応答の測定方法について、実験機材の構成やプログラムなど多大なご助言を頂きました。深く感謝いたします。

龍谷大学理工学部 岩嶋浩樹実験講師には、実環境での音圧測定実験や主観評価実験を行うにあたり、機材や設備について多大なご助言、ご協力をいただきました。また、計算機の使用方法や構成についてご指導をいただきました。深く感謝いたします。

情報通信研究機構 岡本拓磨博士、静岡大学 立蔵洋介准教授には、本研究の内容や理論について学会発表の場で多くのご議論、ご助言をいただきました。ここに深く感謝いたします。また、研究を進めるにあたり、数々の有益なご助言をいただいた日本音響学会 学生・若手フォーラムの諸氏に感謝いたします。

研究中多方向にわたってご協力いただいた龍谷大学理工学部片岡研究室の諸氏に深く感謝いたします。また、研究中多くの励ましをいただいた学内外の友人に深く感謝いたします。

最後になりましたが、今日まで筆者を支え、見守っていただいた家族に感謝いたします。



## 参考文献

- [1] 田村秀行, “複合現実感：現実と仮想の融合,” バイオメカニズム学会誌, Vol. 25, No. 2, pp. 58–61, 2001.
- [2] 山本裕之, “仮想と現実の融合：1. 複合現実感-仮想と現実の境界から見える世界-,” 情報処理, Vol. 43, No. 3, pp. 213–216, 2002.
- [3] 大槻麻衣, 木村朝子, 西浦敬信, 柴田史久, 田村秀行, “複合現実空間との新しいマルチモーダル・インタラクション方法の提案と実現,” 日本バーチャルリアリティ学会論文誌, Vol. 13, No. 2, pp. 247–255, 2008.
- [4] 安藤彰男, “高臨場感音響技術とその理論,” 電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review, Vol. 3, No. 4, pp. 33–46, 2009.
- [5] 安藤彰男, “音場再現,” コロナ社, 2014.
- [6] W. Zhang , “Surround by Sound: A Review of Spatial Audio Recording and Reproduction,” *Applied Sciences*, Vol. 7, No. 5, 2017.
- [7] A. J. Berkhout, D. Vries, and P. Vogel , “Acoustic control by wave field synthesis,” *The Journal of the Acoustical Society of America*, Vol. 93, No. 5, pp. 2764–2778, 1993.
- [8] T. Betlehem and T. D. Abhayapala , “Theory and design of sound field reproduction in reverberant rooms,” *The Journal of the Acoustical Society of America*, Vol. 117, No. 4, pp. 2100–2111, 2005.

- [9] J. Chang, C. Lee, J. Park, and Y. Kim , “A realization of sound focused personal audio system using acoustic contrast control,” *The Journal of the Acoustical Society of America*, Vol. 125, No. 4, pp. 2091–2097, 1983.
- [10] H. Wierstorf, A. Raake, M. Geier, and S. Spors , “Perception of Focused Source in Wave Field Synthesis,” *Journal of the Audio Engineering Society*, Vol. 61, No. 1, pp. 5–16, 2013.
- [11] S. Koyama, K. Furuya, Y. Hiwasaki, Y. Haneda, and Y. Suzuki , “Wave Field Reconstruction Filtering in Cylindrical Harmonic Domain for With-Height Recording and Reproduction,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Vol. 22, No. 10, pp. 1546–1557, Oct. 2014.
- [12] J. W. Choi and Y. H. Kim , “Generation of an acoustically bright zone with an illuminated region using multiple sources,” *The Journal of the Acoustical Society of America*, Vol. 111, No. 4, pp. 1695–1700, Apr. 2002.
- [13] M. Shin, S. Q. Lee, F. M. Fazi, P. A. Nelson, D. Kim, S. Wang, K. H. Park, and J. Seo , “Maximization of acoustic energy difference between two spaces,” *The Journal of the Acoustical Society of America*, Vol. 128, No. 1, pp. 121–131, July 2010.
- [14] 岡本拓磨, “波数領域空間フィルタリングに基づく直線スピーカアレイを用いたマルチスポット再生,” 電子情報通信学会技術研究報告, EA, Vol. 113, No. 349, pp. 37–42, Dec. 2013.
- [15] N. Radmanesh, I. S. Burnett, and B. D. Rao , “A Lasso-LS Optimization with a Frequency Variable Dictionary in a Multizone Sound System,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, Vol. 24, No. 3, pp. 583–593, 2016.

- [16] T. Okamoto and A. Sakaguchi , “Experimental validation of spatial Fourier transform-based multiple sound zone generation with a linear loudspeaker array,” *The Journal of the Acoustical Society of America*, Vol. 141, No. 3, pp. 1769–1780, 2017.
- [17] 野村浩一, 片岡章俊, “多点制御法を用いた目標方向外音声の低明瞭度化手法,” 電子情報通信学会技術研究報告, EA, Vol. 111, No. 402, pp. 13–18, Jan. 2012.
- [18] 安枝和哉, 片岡章俊, “多点制御法を用いた環境音による目的方向外音声のマスキング手法に関する検討,” 電子情報通信学会技術研究報告, EA, Vol. 114, No. 358, pp. 31–36, Dec. 2014.
- [19] K. Yasueda, D. Shinjo, and A. Kataoka , “Arrangement of control points for sound spot generation in the multi-point control method,” *5th Joint Meeting of the ASA and ASJ, The Journal of the Acoustical Society of America*, Vol. 140, No. 4, p. 3140, 2016.
- [20] M. Yoneyama, J. Fujimoto, Y. Kawano, and S. Sasabe , “The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design,” *The Journal of the Acoustical Society of America*, Vol. 73, No. 5, pp. 1532–1536, 1983.
- [21] P. J. Westrevelt , “Parametric acoustic array,” *The Journal of the Acoustical Society of America*, Vol. 35, No. 4, pp. 535–537, 1963.
- [22] 鎌倉友男, 酒井新一, “パラメトリックスピーカの原理と応用,” 電子情報通信学会技術研究報告, US, Vol. 105, No. 554, pp. 19–24, Jan. 2006.
- [23] 小森慎也, 益永翔平, 生藤大典, 中山雅人, 西浦敬信, “フレキシブルパラメトリックスピーカを用いたオーディオスポットの制御,” 日本バーチャルリアリティ学会論文誌, Vol. 20, No. 3, pp. 189–198, Mar. 2015.

- [24] 鎌倉友男, 伊藤幹也, 野村英之, “超音波暴露の調査: パラメトリックスピーカを話題にして,” 日本音響学会誌, Vol. 67, No. 5, pp. 200–203, May 2011.
- [25] M. M. Boone, W. H. Cho, and J. G. Ih , “Design of a highly directional endfire loudspeaker array,” *Journal of the Audio Engineering Society*, Vol. 57, No. 5, pp. 309–325, 2009.
- [26] 西川清, 大崎貴也, “2次元デジタルフィルタを用いた指向性アレイスピーカ,” 電子情報通信学会論文誌 (A), Vol. 78, No. 11, pp. 1419–1428, Nov. 1995.
- [27] 羽田陽一, “音の波数領域信号処理,” 電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review, Vol. 11, No. 4, pp. 243–255, 2018.
- [28] 岡本拓磨, “角度スペクトルから円調和関数展開への解析的音場変換,” 日本音響学会講演論文集 (春), pp. 763–766, Mar. 2016.
- [29] T. Okamoto , “2.5D higher-order Ambisonics for a sound field described by angular spectrum coefficients,” *Proceedings of ICASSP 2016*, pp. 326–330, Mar. 2016.
- [30] 関貴志, 羽田陽一, “音場制御と聴覚マスキングの併用による円形スピーカアレイを用いたエリア再生法の提案,” 日本音響学会講演論文集 (春), pp. 851–854, 2016.
- [31] O. Kirkeby and P. A. Nelson , “Reproduction of plane wave sound fields,” *The Journal of the Acoustical Society of America*, Vol. 94, No. 5, pp. 2992–3000, Nov. 1993.
- [32] O. Kirkeby, P. A. Nelson, F. Orduna-Bustamante, and H. Hamada , “Local sound field reproduction using digital signal processing,” *The Journal of the Acoustical Society of America*, Vol. 100, No. 3, pp. 1584–1593, Nov. 1996.
- [33] P. A. Gauthier and A. Barry , “Sound-field reproduction in-room using optimal techniques: Simulations in the frequency domain,” *The Journal of the Acoustical Society of America*, Vol. 117, No. 2, pp. 662–678, 2005.

- [34] S. Miyabe, H. Saruwatari, K. Shikano, and Y. Tatekura , “Interface for Barge-in Free Spoken Dialogue System Using Nullspace Based Sound Field Control and Beamforming,” *IEICE transactions on fundamentals of electronics, communications and computer sciences*, Vol. 89, No. 3, pp. 716–726, 2006.
- [35] J. Ahrens and S. Spors , “Sound field reproduction using planar and linear arrays of loudspeakers,” *IEEE/ACM Transactions on Audio Speech and Language Processing*, Vol. 18, No. 8, pp. 2038–2050, 2010.
- [36] E. G. Williams , “Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography,” *Academic Press*, 1999.
- [37] 神沼充伸, 伊勢史郎, 鹿野清宏 , “周波数領域における最小ノルム解を利用した多チャンネル音場再現システムにおける逆フィルタの設計,” *日本音響学会誌*, Vol. 57, No. 3, pp. 175–183, 2001.
- [38] M. Born and E. Wolf , “Principles of Optics,” *Cambridge University Press*, 1999.
- [39] ISO 10534-2 , “Acoustics - Determination of sound absorption coefficient and impedance in impedance tubes - Part2: Transfer-function method,” 1998.
- [40] 伊勢史郎, 鮫島俊哉, 山崎芳男 , “キルヒホッフ積分 方程式に基づく音場制御理論とその応用,” *建築音響研究会*, Vol. AA94, No. 5, 1994.
- [41] 伊勢史郎, “キルヒホッフ-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*, Vol. 53, No. 9, pp. 706–713, Sep. 1997.
- [42] 榎本成悟, 伊勢史郎, “境界音場制御の原理を用いた指向性スピーカシステムの提案,” *電子情報通信学会論文誌 (A)*, Vol. 87, No. 4, pp. 431–438, Apr. 2004.
- [43] 橋本泰治, 三河正彦, 田中和世 , “境界音場制御の原理を用いた全方位音響マルチビームフォーミングシステムの開発,” *日本音響学会誌*, Vol. 68, No. 3, pp. 113–120, Mar. 2012.

- [44] H. Tokuno, O. Kirkeby, P. A. Nelson, and H. Hamada , “Inverse Filter of Sound Reproduction Systems Using Regularization,” *IEICE transactions on fundamentals of electronics, communications and computer sciences*, Vol. 80, No. 5, pp. 809–820, May 1997.
- [45] 植松尚, 羽田陽一, 片岡章俊, “エリア内の再生特性を考慮した音の局所的再生,” 日本音響学会誌, Vol. 62, No. 2, pp. 89–97, Feb. 2006.
- [46] 羽田陽一, 片岡章俊, “自由空間伝達関数を用いた多点制御に基づく小型スピーカアレーの実空間性能,” 日本音響学会講演論文集 (春), pp. 631–632, Mar. 2008.
- [47] 坂本修一, 鈴木陽一, 天野成昭, 小澤賢司, 近藤公久, 曾根敏夫, “親密度と音韻バランスを考慮した単語了解度試験用リストの構築,” 日本音響学会誌, Vol. 54, No. 12, pp. 842–849, Dec. 1998.
- [48] ITU-T P.800 , “Methods for subjective determination of transmission quality,” 1996.
- [49] 根木教男, 及川靖広, 服部永雄, 山崎芳男, “ランダムベクトル信号の空間合成による音響信号の再現,” 日本音響学会講演論文集 (春), pp. 497–498, Mar. 2000.
- [50] 高根昭一, 安部幸治, 渡邊貫治, 佐藤宗純, “信号分解に基づく音声のスポット再生手法に関する一考察,” 日本音響学会講演論文集 (秋), pp. 675–676, Sep. 2010.
- [51] 戸倉綾, 三浦治, 東山三樹夫, “音声信号のランダムベクトル表現,” 日本音響学会講演論文集 (秋), pp. 497–498, Sep. 1999.
- [52] 杉耕作, 及川靖広, 山崎芳男, “ランダムベクトル信号による音響信号展開方法の検討,” 日本音響学会講演論文集 (秋), pp. 717–718, 2001.
- [53] 佐藤洋, 清水寧, “スピーチプライバシー研究の歴史と近年の動向,” 日本音響学会誌, Vol. 64, No. 8, pp. 475–480, Aug. 2008.

- [54] 赤木正人, 入江佳洋, “音情景理解を応用した音声プライバシー保護,” 電子情報通信学会技術研究報告, EA, Vol. 111, No. 333, pp. 19–24, Dec. 2011.
- [55] 佐伯徹郎, 藤井健生, 山口静馬, 老松建成, “音声をマスクするための無意味定常雑音の選定,” 電子情報通信学会論文誌 (A), Vol. 86, No. 2, pp. 187–191, Feb. 2003.
- [56] 佐伯徹郎, 山口静馬, 為末隆弘, “マスキングノイズによるスピーチプライバシー保護に関する一考察,” 日本音響学会論文誌, Vol. 61, No. 10, pp. 571–575, 2005.
- [57] 山川高史, 秦雅人, 藤原舞, 清水寧, “薬局におけるカウンターと待合席間のスピーチプライバシーソリューションの一提案,” 日本音響学会講演論文集 (秋), pp. 1123–1126, Sep. 2011.
- [58] 佐藤逸人, “日本におけるスピーチプライバシー,” 騒音制御, Vol. 33, No. 4, pp. 287–290, Aug. 2009.
- [59] A. Ito, A. Miki, Y. Shimizu, K. Ueno, H. J. Lee, and S. Sakamoto, “Oral information masking considering room environmental condition, Part 1 : Synthesis of Maskers and examination on their masking efficiency,” *Proceedings of Inter-Noise 2007*, 2007.
- [60] ITU-T P.862, “Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs,” 2002.
- [61] 藤本卓也, “低域バンドでの SN 比改善を目的とした TSP 信号に関する検討,” 日本音響学会講演論文集 (秋), pp. 432–433, Sep. 1999.
- [62] D. E. Goldberg, “Genetic Algorithms in Search, Optimization and Machine Learning,” *Addison-Wesley*, 1989.
- [63] 伊庭齊志, “遺伝的アルゴリズムの基礎 GA の謎を解く,” コロナ社, 1994.

- [64] 久保幹雄, J. P. ペドロソ, “メタヒューリスティクスの数理,” 共立出版, 2009.
- [65] 伊藤弘章, 古家賢一, 羽田陽一, “平面スピーカアレイを用いたエバネッセント波再生手法,” 電子情報通信学会技術研究報告, EA, Vol. 110, No. 471, pp. 29–34, Mar. 2011.
- [66] 岡本拓磨, “平面複数円形アレイを用いた3次元近傍エリア再生,” 日本音響学会講演論文集 (秋), pp. 195–198, Sep. 2019.

# 研究業績

## 学術論文

1. 安枝和哉, 野村浩一, 片岡章俊, “多点制御法と音声の残差信号分解法を組み合わせた音のエリア再生法,” 電子情報通信学会論文誌 A, Vol.J97-A, No.2, pp.77-85, 2014
2. 安枝和哉, 大野瑞季, 片岡章俊, “目的方向外での会話の漏えいを防ぐ多点制御法による環境音再生法,” 電子情報通信学会論文誌 A, Vol.J98-A, No.11, pp.610-618, 2015
3. 安枝和哉, 新城大輔, 片岡章俊, “直線スピーカアレーを用いた多点制御法による複数エリア再生,” 電子情報通信学会論文誌 A, Vol.J103-A, No.1, pp.9-16, 2020.

## 国際会議

1. **Kazuya YASUEDA**, Daisuke SHINJO, Akitoshi KATAOKA, ”Arrangement of control points for sound spot generation in the multi-point control method,” *5th Joint Meeting of the Acoustical Society of America and Acoustical Society of Japan, The Journal of the Acoustical Society of America*, Vol. 140, No. 4, p.3140, Oct. 2016.
2. **Kazuya YASUEDA**, Daisuke SHINJO, Akitoshi KATAOKA, ”Determination of Optimal Parameters Using Metaheuristics for the Sound Zone Generation by the Least-Squares,” *The proceedings of the ICA2019*, pp.2708-2715, Sep. 2019.

## 研究会

1. 安枝和哉, 片岡章俊, “多点制御法と音声の残差信号分解法を用いた音のエリア再生,” 電子情報通信学会技術報告, vol. 112, no. 478, EA2012-147, pp. 37-42, Mar. 2013.
2. 安枝和哉, 片岡章俊, “多点制御法を用いた環境音による目的方向外音声のマスキング手法に関する検討,” 電子情報通信学会技術報告, vol.114, no. 358, EA2014-45, pp.31-36, Dec. 2014.
3. 安枝和哉, 新城大輔, 片岡章俊, “実環境における多点制御法を用いた複数エリア再生,” 電子情報通信学会技術報告, vol. 116, no. 475, EA2016-143, pp.339-344, Mar. 2017.
4. 安枝和哉, 新城大輔, 片岡章俊, “複数スピーカアレーを用いた分解音声の局所的再合成によるスポット再生,” 電子情報通信学会技術報告, vol. 117, no. 515, EA2017-119, pp.107-110, Mar. 2018.
5. 新城大輔, 安枝和哉, 片岡章俊, “多点制御法を用いた複数エリア再生における遺伝的アルゴリズムを活用した制御点配置の決定方法,” 電子情報通信学会技術報告, vol. 118, no. 357, EA2018-83, pp.15-20, Dec. 2018.

## 国内大会・研究交流会

1. 安枝和哉, 片岡章俊, “多点制御法における目的方向外音声の環境音によるマスキングに関する検討,” 日本音響学会春季研究発表会, pp.771-772, Mar. 2014.
2. 安枝和哉, 冠野美利, 片岡章俊, “多点制御された環境音における方向別の再生品質に関する検討,” 日本音響学会秋季研究発表会, pp.837-838, Sep. 2014.

3. 安枝和哉, 片岡章俊, “エリア再生における信号分解法とマスキング手法を用いたエリア外での秘匿性向上の検討,” 第17回日本音響学会関西支部若手研究者交流研究発表会, 32, Dec. 2014.
4. 安枝和哉, 片岡章俊, “多点制御法における信号分解法とマスキング手法を用いた目的方向外での秘匿性向上,” 日本音響学会春季研究発表会, pp.773-774, Mar. 2015.
5. 安枝和哉, 新城大輔, 片岡章俊, “多点制御法における格子状制御点配置によるマルチスポット再生,” 日本音響学会秋季研究発表会, pp.651-652, Sep. 2016.
6. 安枝和哉, 新城大輔, 片岡章俊, “ランダムベクトル分解信号を用いた局所再現に関する検討,” 第19回日本音響学会関西支部若手研究者交流研究発表会, 49, Dec. 2016.
7. 安枝和哉, 新城大輔, 片岡章俊, “複数信号の局所的再合成によるエリア再生の検討,” 日本音響学会春季研究発表会, pp.611-612, Mar. 2017.
8. 安枝和哉, 新城大輔, 片岡章俊, “複数スピーカアレーを用いた分解信号の再合成によるエリア再生,” 日本音響学会秋季研究発表会, pp.739-740, Sep. 2017.
9. 安枝和哉, 片岡章俊, “多点制御法を用いた音のエリア/スポット再生に関する研究,” 日本音響学会 学生研究交流会, Mar. 2018.
10. 安枝和哉, 新城大輔, 片岡章俊, “仮想焦点音源を用いた多点制御法による分解音声の局所的再合成,” 日本音響学会春季研究発表会, pp.517-518, Mar. 2018.
11. 新城大輔, 安枝和哉, 片岡章俊, “エリア再生のための非負値行列因子分解を用いた信号分解法,” 日本音響学会春季研究発表会, pp.519-520, Mar. 2018.
12. 齊藤由依, 安枝和哉, 吉水琢人, 片岡章俊, “音場推定に基づくエリア收音における位置推定誤差に関する検討,” 日本音響学会春季研究発表会, pp.543-544, Mar. 2018.

13. 安枝和哉, 片岡章俊, “スピーチプライバシー保護のための境界音場制御を用いた環境音再生手法の検討,” 日本音響学会秋季研究発表会, pp.249-250, Sep. 2018.
14. 齊藤由依, 安枝和哉, 片岡章俊, “複数マイクロホンアレイを用いた音源の位置関係に基づくエリア收音,” 日本音響学会秋季研究発表会, pp.417-418, Sep. 2018.
15. 安枝和哉, 新城大輔, 高岡克典, 片岡章俊, “多点制御法による複数エリア再生システムの実装と評価,” 第21回日本音響学会関西支部若手研究者交流研究発表会, 28, Dec. 2018.
16. 齊藤由依, 安枝和哉, 大下智葵, 片岡章俊, “遅延和ビームフォーマの指向特性に基づくサブアレイ選択による音声強調,” 日本音響学会春季研究発表会, pp.325-326, Mar. 2019
17. 安枝和哉, 新城大輔, 高岡克典, 片岡章俊, “多点制御法を用いた局所再生における抑圧エリアの形成に関する検討,” 日本音響学会春季研究発表会, pp.327-328, Mar. 2019.
18. 安枝和哉, 片岡章俊, “音質を考慮した評価関数に基づく多点制御法の制御点配置決定法,” 日本音響学会秋季研究発表会, pp.663-666, Sep. 2019.
19. 野浪初音, 安枝和哉, 片岡章俊, “U-Net 構造を持った CNN による雑音抑圧の検討,” 第22回日本音響学会関西支部若手研究者交流研究発表会, 33, Dec. 2019.

