

2次元デジタルフィルタを用いる 指向性制御方法の応用例

金沢大学 オーディオ情報処理研究室
伊藤 栄太, 山川 昭彦, 齋藤 毅, 三好 正人

目次

- ①マイクロホンアレーによる移動音声の高S/N收音について
- ②集束ビームを用いて生成した仮想音源の評価

マイクロホンアレーによる移動音声の高S/N收音について

1.はじめに

- 講演者が移動しながら話をする時

空調雑音



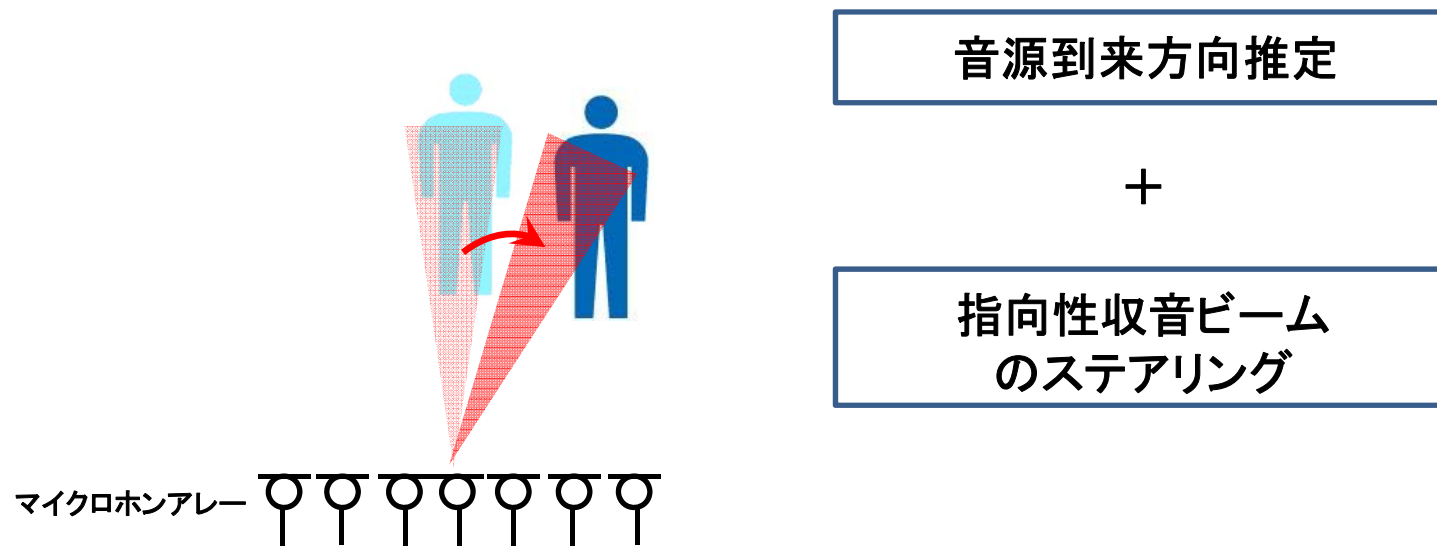
生徒のざわめき

会場内に設置した固定マイクロホンで
講演者音声のみを收音したい

2.提案システム

2.1 システム原理

- 固定マイクロホンで、講演者音声のみを高S/N收音する



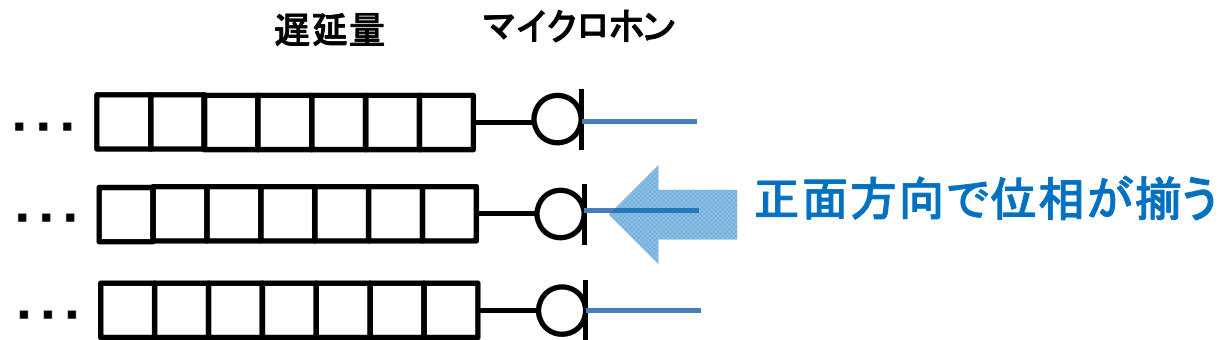
推定方向に指向性收音ビームをステアリングする

2.提案システム

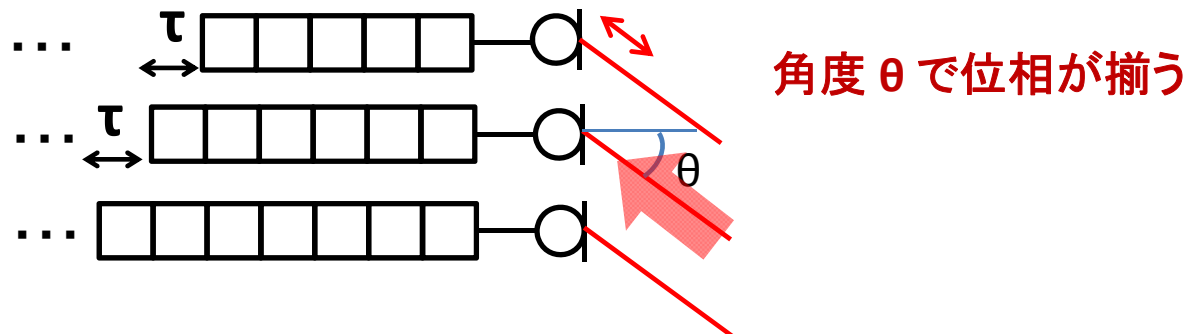
2.1 システム原理

- ビームステアリング

- ・全てのマイクにかける遅延量が同じ場合



- ・マイク毎の遅延量を変化させる場合

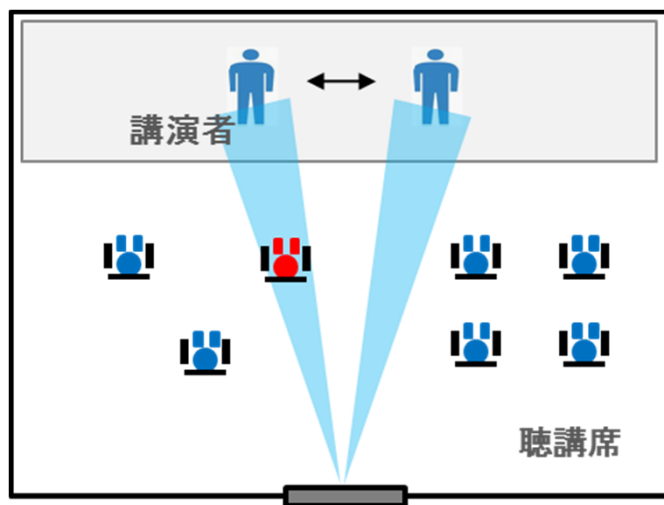


2.提案システム

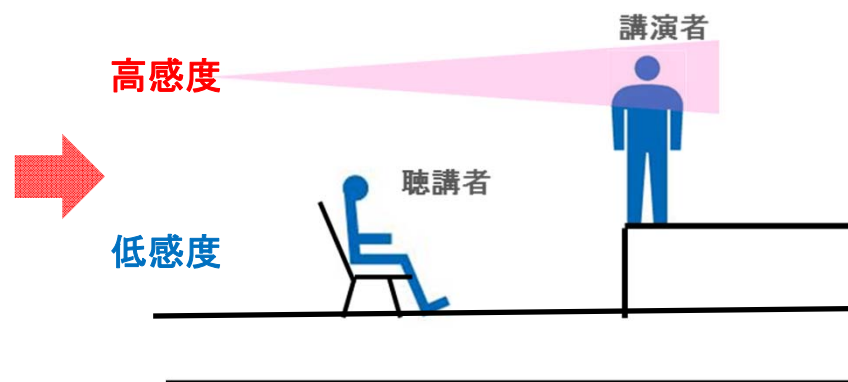
2.1 システム原理

- ビーム内に雑音を含んでしまう場合

左右方向が分かったとしても...



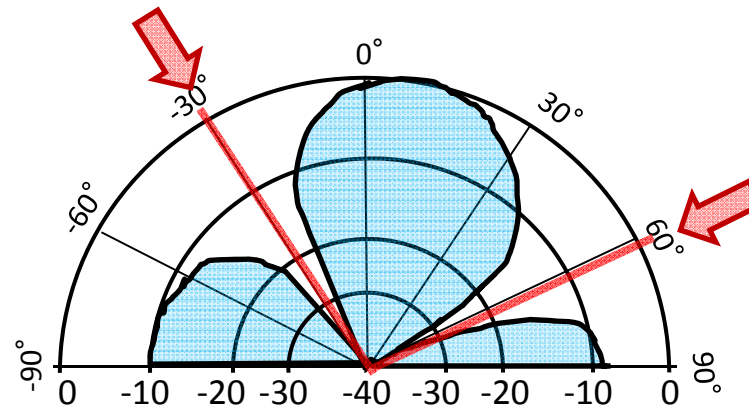
聴講者音声などが含まれる場合がある



音源の高さの違いを利用する

2.提案システム

2.2 ヌルビームによる方向推定

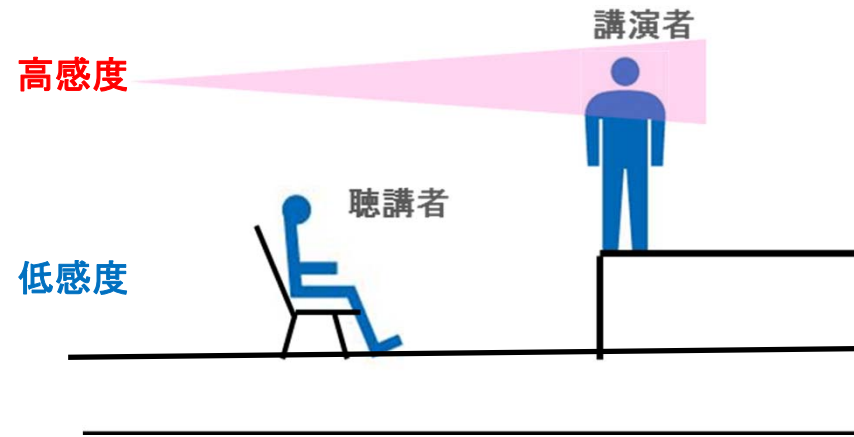


ヌルビームの問題点

- ①複数の方向のうち、どれが講演者の音声なのかを判断する基準が必要
- ②方向推定出来る音源数が「アレーの総素子数-1」個に限られる
※以下、アレーの自由度と呼ぶ

2.提案システム

2.2 ヌルビームによる方向推定



解決案

①複数の方向のうちどれが講演者の音声なのかを判断する基準が必要 ⇒パワーの最も大きい方向に定める

②方向推定出来る音源数が「アレーの総素子数-1」個に限られる

※自由度

⇒パワーの大きい到来方向を優先させる

2.提案システム

2.2 ヌルビームによる方向推定

<ヌルビームによる 到来方向推定法 >

超分解能

指導原理	素子数(K)と到来波数(L)	計算原理	特徴	備考
線形予測法	$K-1=L$	方向サーチ	予測誤差最小化	熱雑音の影響有
最小ノルム法	$K-1=L$	固有値, 方向サーチ	電力最小化	-
MUSIC法	$K-1 \geq L$	固有値, 方向サーチ	電力最小化	-
ESPRIT法	$K-1 \geq L$	固有値	サブアレー間位相差	アレイ形状は同形

音源数が自由度を越えてもパワーの大きい方向を優先させる, **最小ノルム法**を採用

2.提案システム

2.3 最小ノルム法の原理

- アレーの出力電力を最小化する方向を求める
⇒それぞれのマイクの出力信号間で相関行列を作成し、その行列の固有値・固有ベクトルを利用する

最小ノルム法の計算式

$$\min(P_{out} = \frac{1}{2} \mathbf{W}^H R_{xx} \mathbf{W})$$

subject to $\mathbf{W}^H \mathbf{W} = 1$

$$P_{MN}(\theta) = \frac{1}{|\mathbf{W}_{MN}^H a(\theta)|^2}$$

(\mathbf{W}_{MN} : R_{xx} の最小固有値の固有ベクトル)

$P_{MN}(\theta)$ が高い値を取る θ が到来方向となる

P_{out} :アレーの出力電力

R_{xx} :マイク間の相関行列

\mathbf{W} :素子ウェイト

2.提案システム

2.3 最小ノルム法の原理

Lagrangeの未定係数法により置き換える

$$Q(\mathbf{W}) = \frac{1}{2} \mathbf{W}^H R_{xx} \mathbf{W} + \frac{\lambda}{2} (1 - \mathbf{W}^H R_{xx} \mathbf{W})$$

λ : 未定係数



$$\nabla_{\mathbf{W}} Q(\mathbf{W}) = \mathbf{R}_{xx} \mathbf{W} - \lambda \mathbf{W} = 0$$



$$\therefore \underline{R_{xx} \mathbf{W} = \lambda \mathbf{W}}$$

固有値問題

2.提案システム

2.3 最小ノルム法の原理

$R_{xx} \mathbf{W} = \lambda \mathbf{W}$ の両辺に \mathbf{W}^H を掛けると

$$\mathbf{W}^H R_{xx} \mathbf{W} = \lambda \mathbf{W}^H \mathbf{W} = \lambda$$

($P_{out} = \frac{1}{2} \mathbf{W}^H R_{xx} \mathbf{W}$) より, λ は出力電力の2倍に相当することが分かる

最小の λ に対応する固有ベクトル \mathbf{W} を用いて,以下の式を得る

$$P_{MN}(\theta) = \frac{1}{|\mathbf{W}_{MN}^H a(\theta)|^2}$$

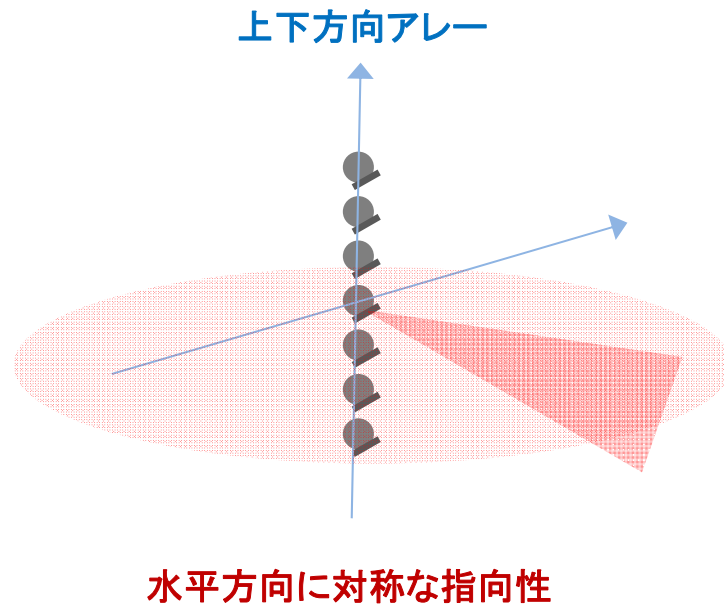
(\mathbf{W}_{MN} : R_{xx} の最小固有値の固有ベクトル)

$P_{MN}(\theta)$ が高い値を取る θ が到来方向となる

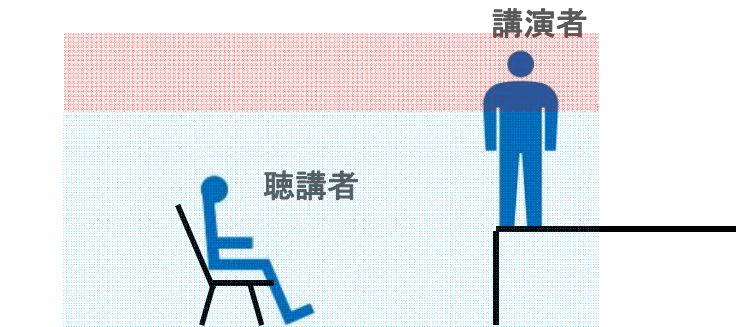
2.提案システム

2.3 指向性制御

- 講演者音声の感度が高い指向性の実現



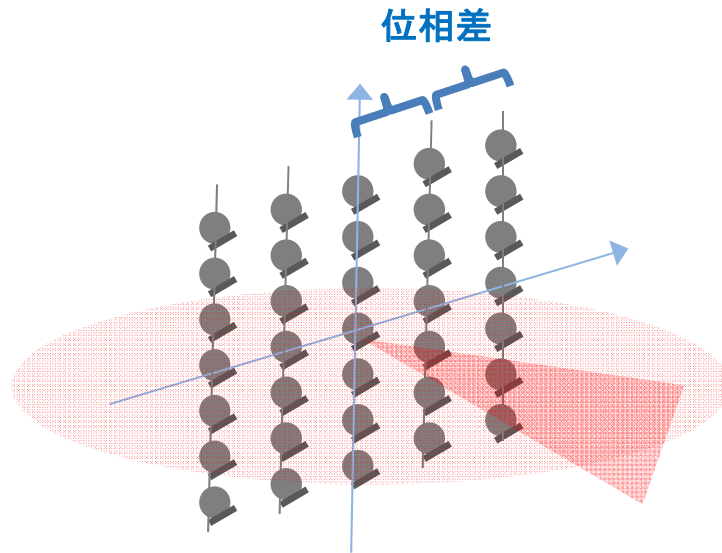
1つの高さ方向アレーを【1ユニット】とする



2.提案システム

2.3 指向性制御

- 話者の左右方向移動での方向推定



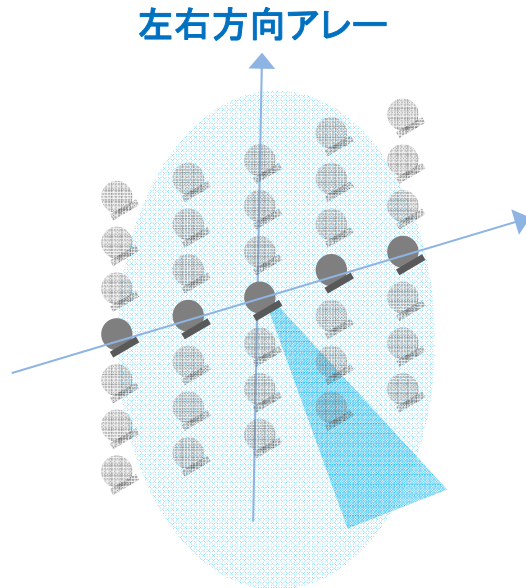
水平に並べた複数の高さユニット間で
左右方向の方向推定

2.提案システム

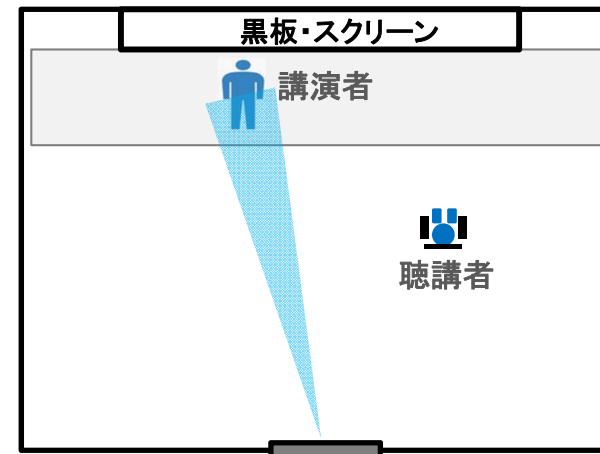
2.3 指向性制御

- 左右方向に指向性收音ビームを向ける

高さアレーの中心素子を代表させて...



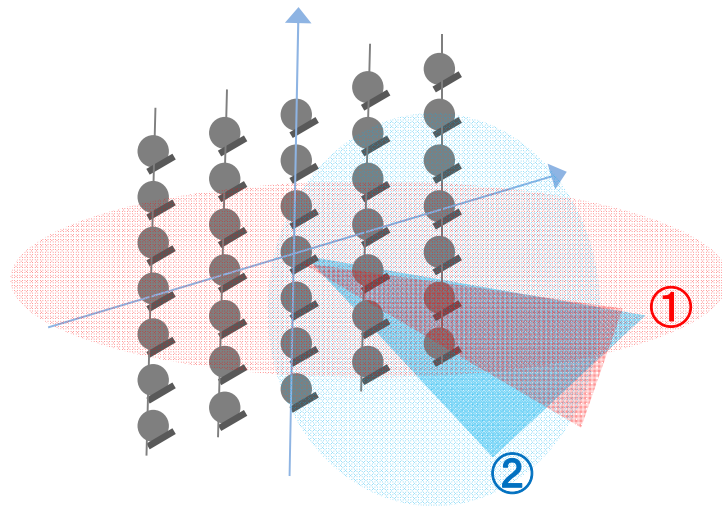
鉛直方向に対称な指向性



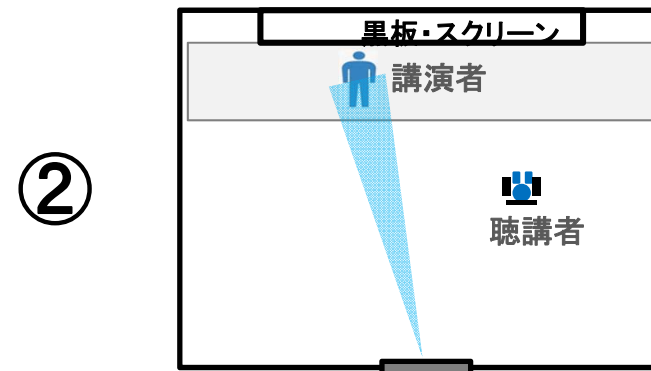
2.提案システム

2.3 指向性制御

- 推定された話者方向に指向性收音ビームを向ける



水平・鉛直指向性の積となり
スポットライト型ビームで講演者音声が強調される

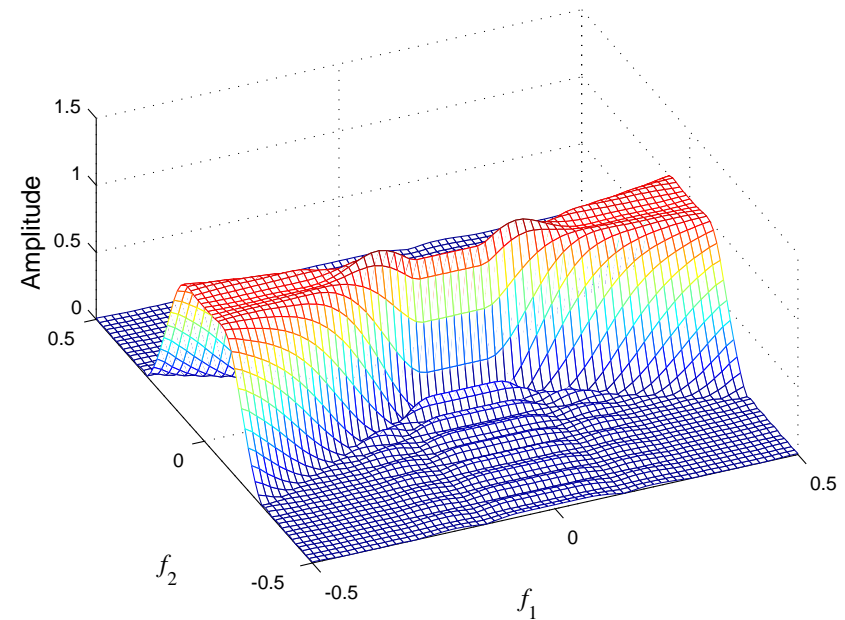


2.提案システム

2.4 指向性ビームの設計

マイク間隔	マイク数	フィルタ次数
4.25 [cm]	19	60
サンプリング周波数		
8000 [Hz]		

周波数	通過域幅
400 [Hz]	$\pm 20^\circ$
周波数	通過域幅
3600 [Hz]	$\pm 12^\circ$



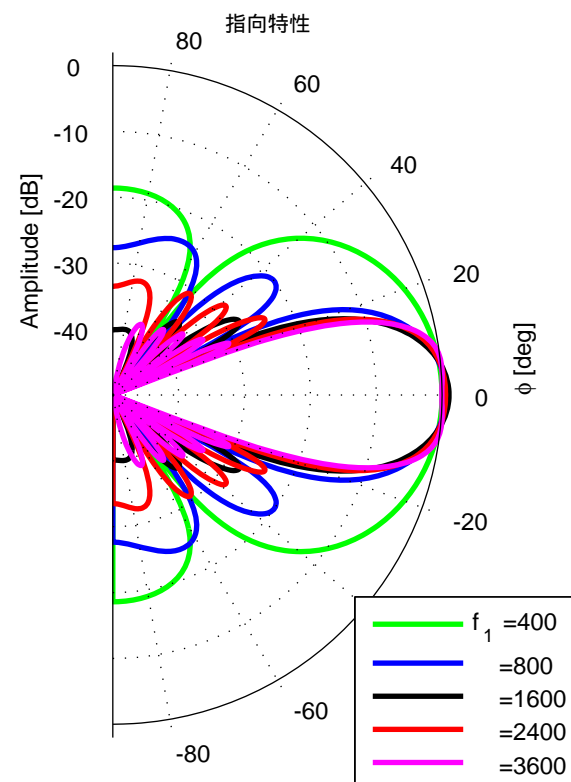
< 振幅特性 >

2.提案システム

2.4 指向性ビームの設計

マイク間隔	マイク数	フィルタ次数
4.25 [cm]	19	60
サンプリング周波数		
8000 [Hz]		

周波数	通過域幅
400 [Hz]	$\pm 20^\circ$
周波数	通過域幅
3600 [Hz]	$\pm 12^\circ$



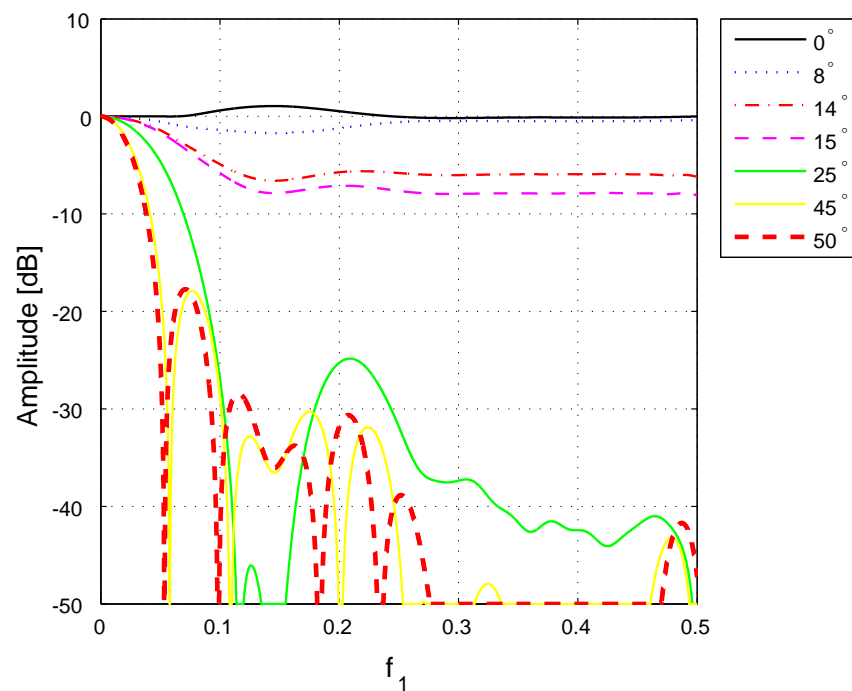
< 指向特性 >

2.提案システム

2.4 指向性ビームの設計

マイク間隔	マイク数	フィルタ次数
4.25 [cm]	19	60
サンプリング周波数		
8000 [Hz]		

周波数	通過域幅
400 [Hz]	$\pm 20^\circ$
周波数	通過域幅
3600 [Hz]	$\pm 12^\circ$

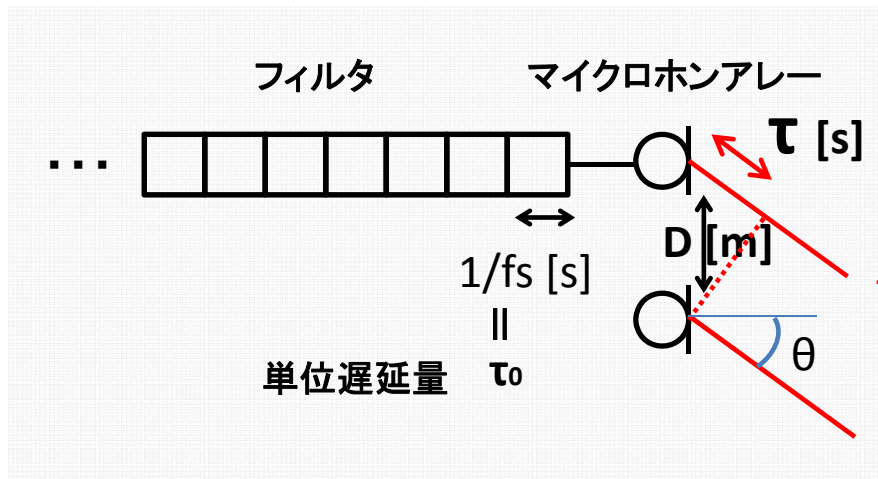


< 周波数特性 >

3. シミュレーション

3.1 条件定義

・受信時のサンプリング周波数の決定



$$\tau = \frac{D \cdot \sin \theta}{c}$$

単位遅延量に合わせて

$$\text{測定できる角度} : \theta = \sin^{-1} \left(\frac{c \cdot (\tau_0 \cdot n)}{D} \right)$$

< 測定可能角度 >

例えば	サンプリング周波数	40[kHz]	のとき...	単位遅延数n	1	2	3	4
	間隔D	4.25[cm]		θ	11.5	23.6	36.9	53.2

3.シミュレーション

3.1 条件定義

・受信時のサンプリング周波数の決定

今回は10° 刻みの移動を追う

< 推定可能な角度分解能 >

サンプリング周波数	間隔D	単位遅延数	1	2	3	4	5	6	7
64[kHz]	4.25[cm]	対応角度	7	14	22	30	39	49	61

サンプリング周波数	64 [kHz]
間隔d	4.25 [cm]

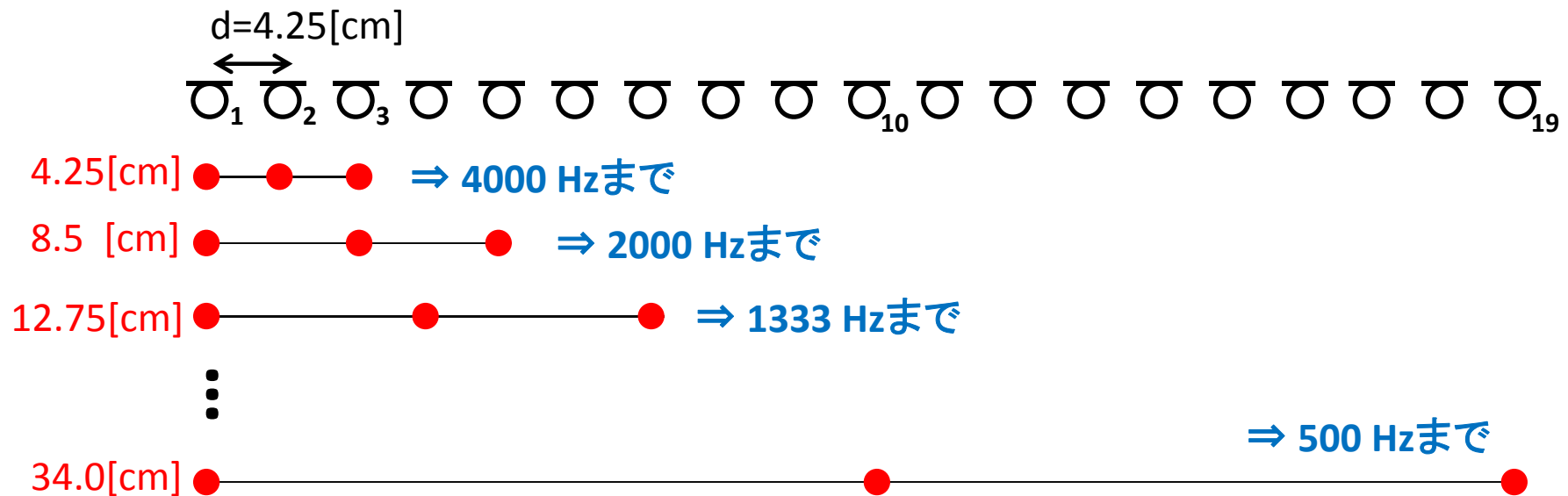
を採用する

3.シミュレーション

3.1 シミュレーション方法

対象周波数 (1/3オクターブバンド中心周波数)

500Hz,600Hz,800Hz,1000Hz,1200Hz,1600Hz,2000Hz,2500Hz,3000Hz

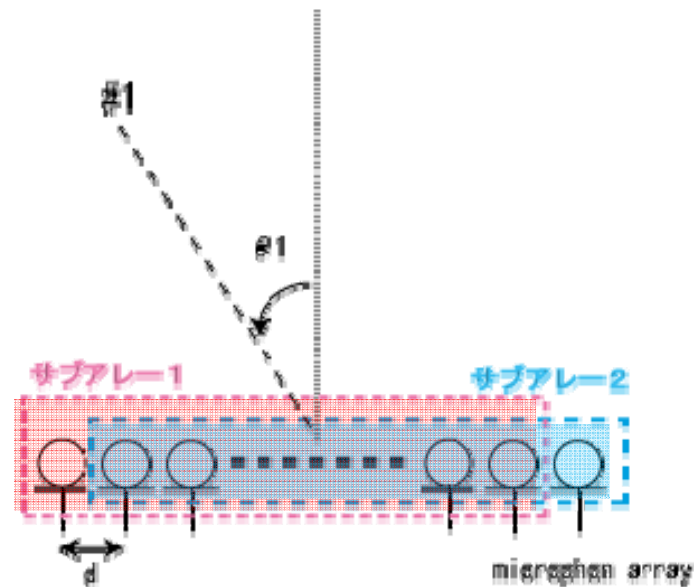


それぞれの周波数の半波長に近い間隔の、**3個のユニット**を適宜選ぶ

3. シミュレーション

3.1 条件定義

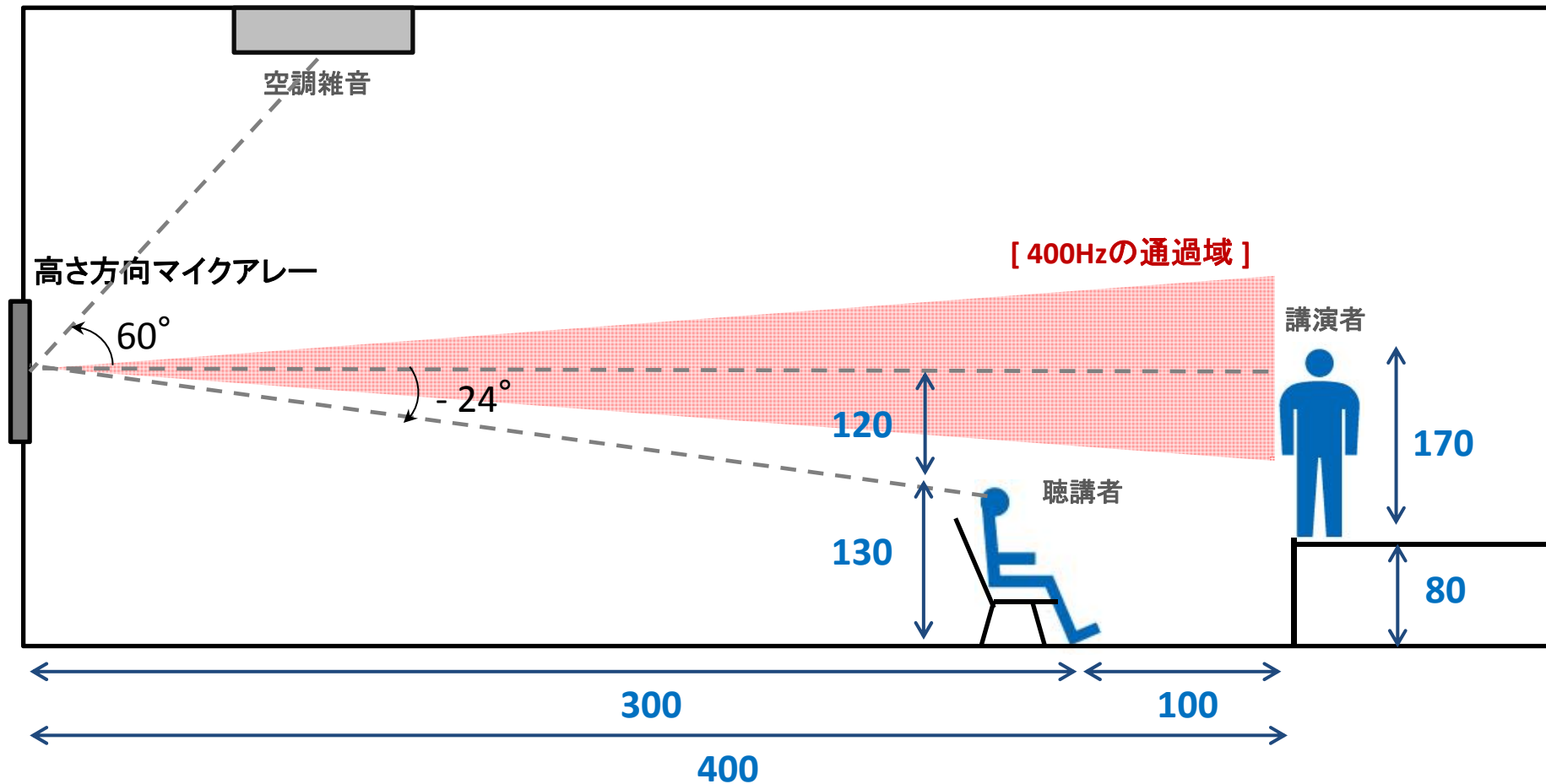
- ESPRIT法・・・2つのサブアレー間の位相差を固有値計算により求めて方向を得る



- 推定精度が非常に高い
- × 音源数が自由度を越えると推定できなくなる

3.シミュレーション

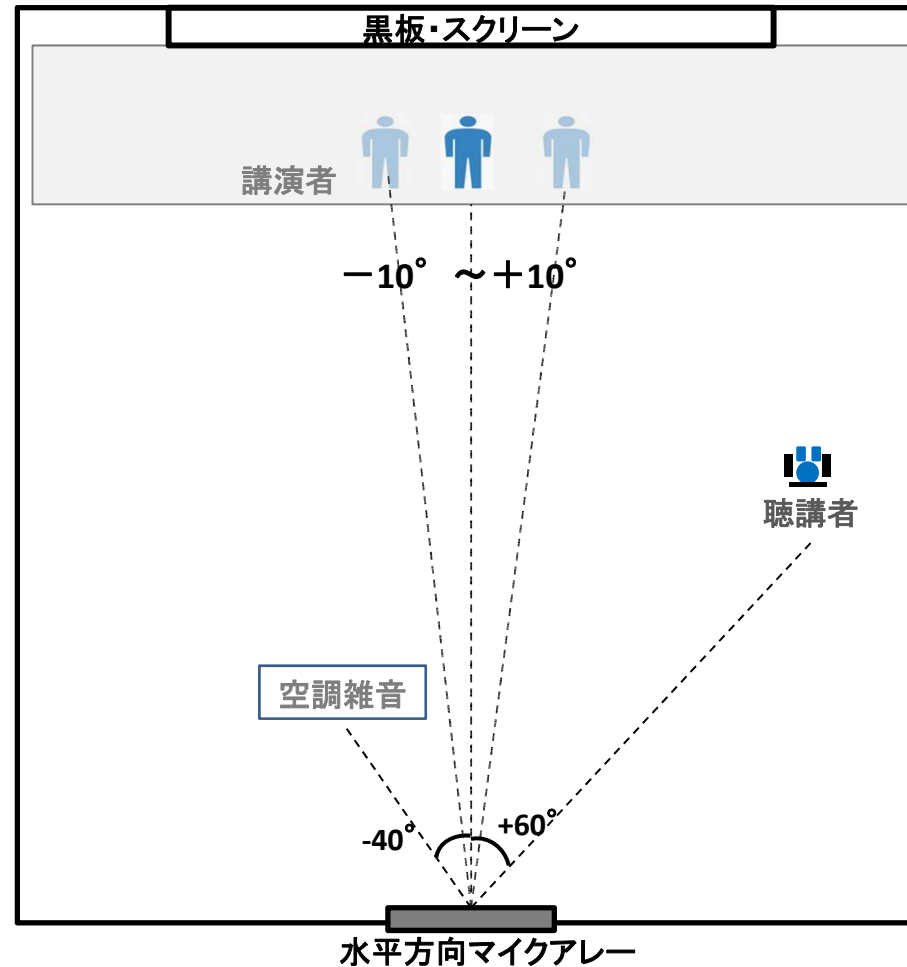
3.2 シミュレーション方法



< 高さ方向で見た部屋の空間配置 >

3. シミュレーション

3.2 シミュレーション方法

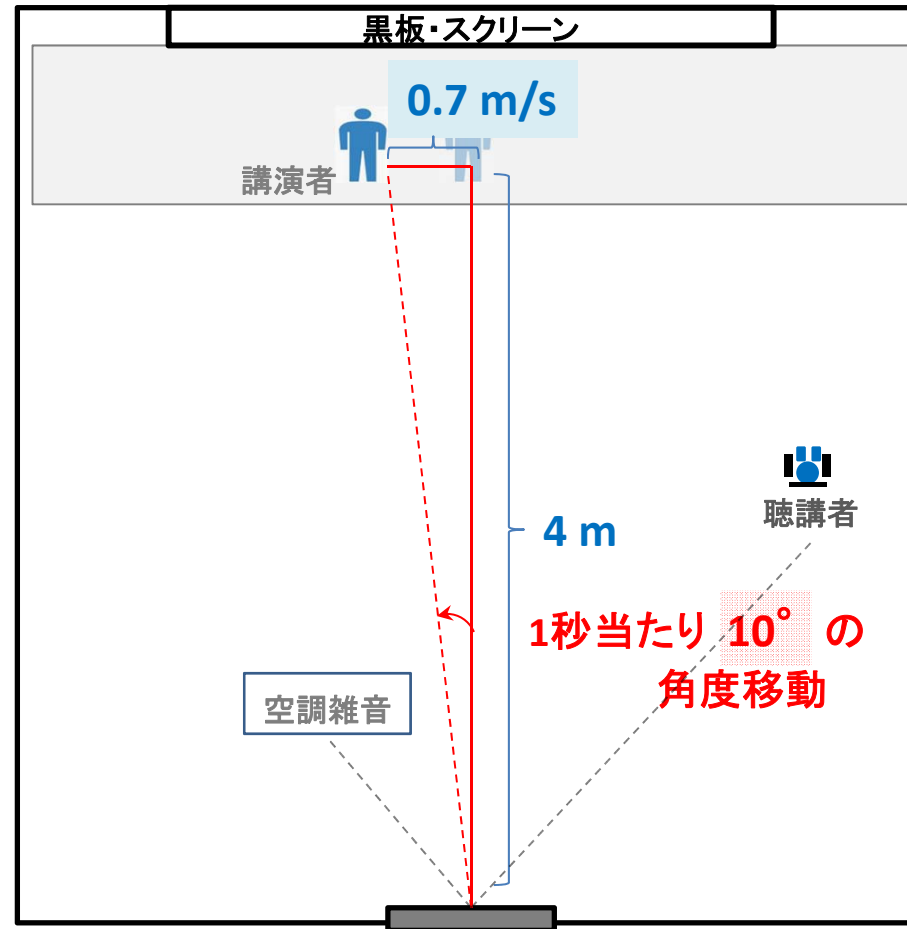


< 水平方向で見た部屋の空間配置 >

3. シミュレーション

3.2 シミュレーション方法

講演者の移動速度を
0.7 m/s とすると...

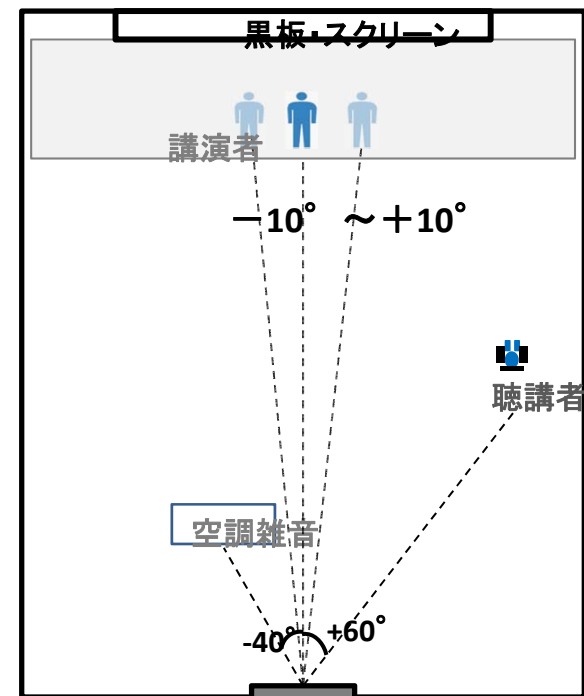
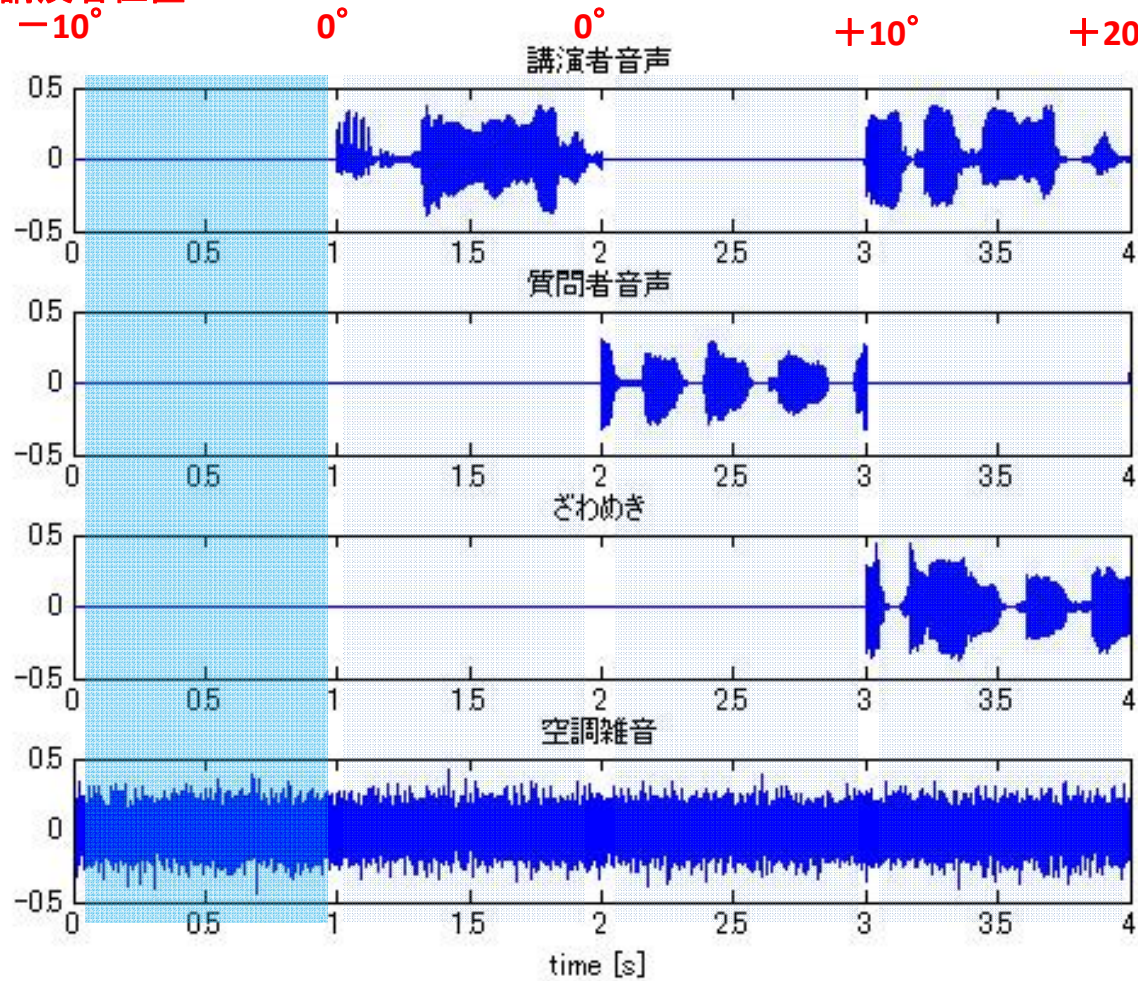


< 水平方向で見た部屋の空間配置 >

3.シミュレーション

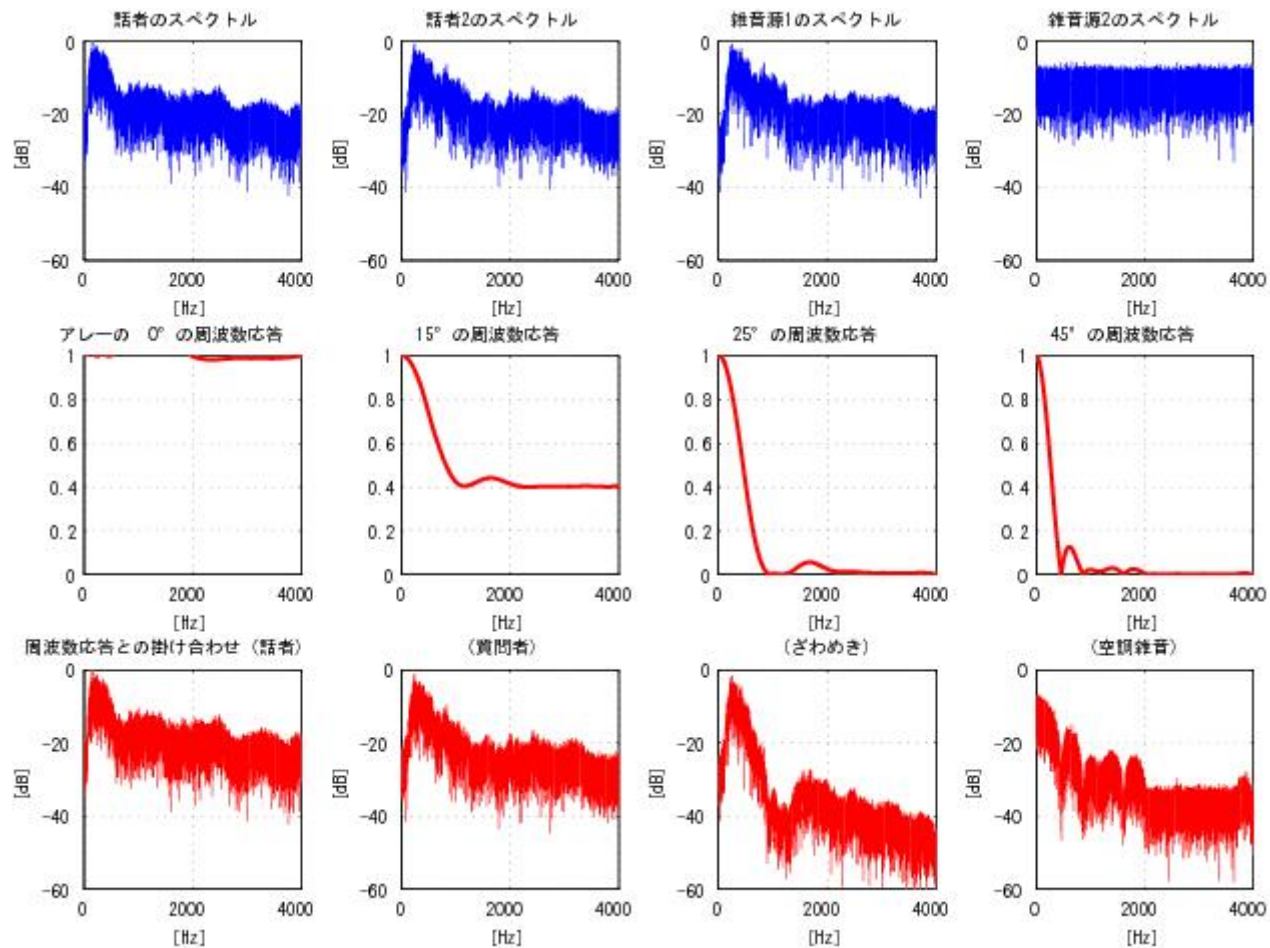
3.2 シミュレーション方法

講演者位置
-10°



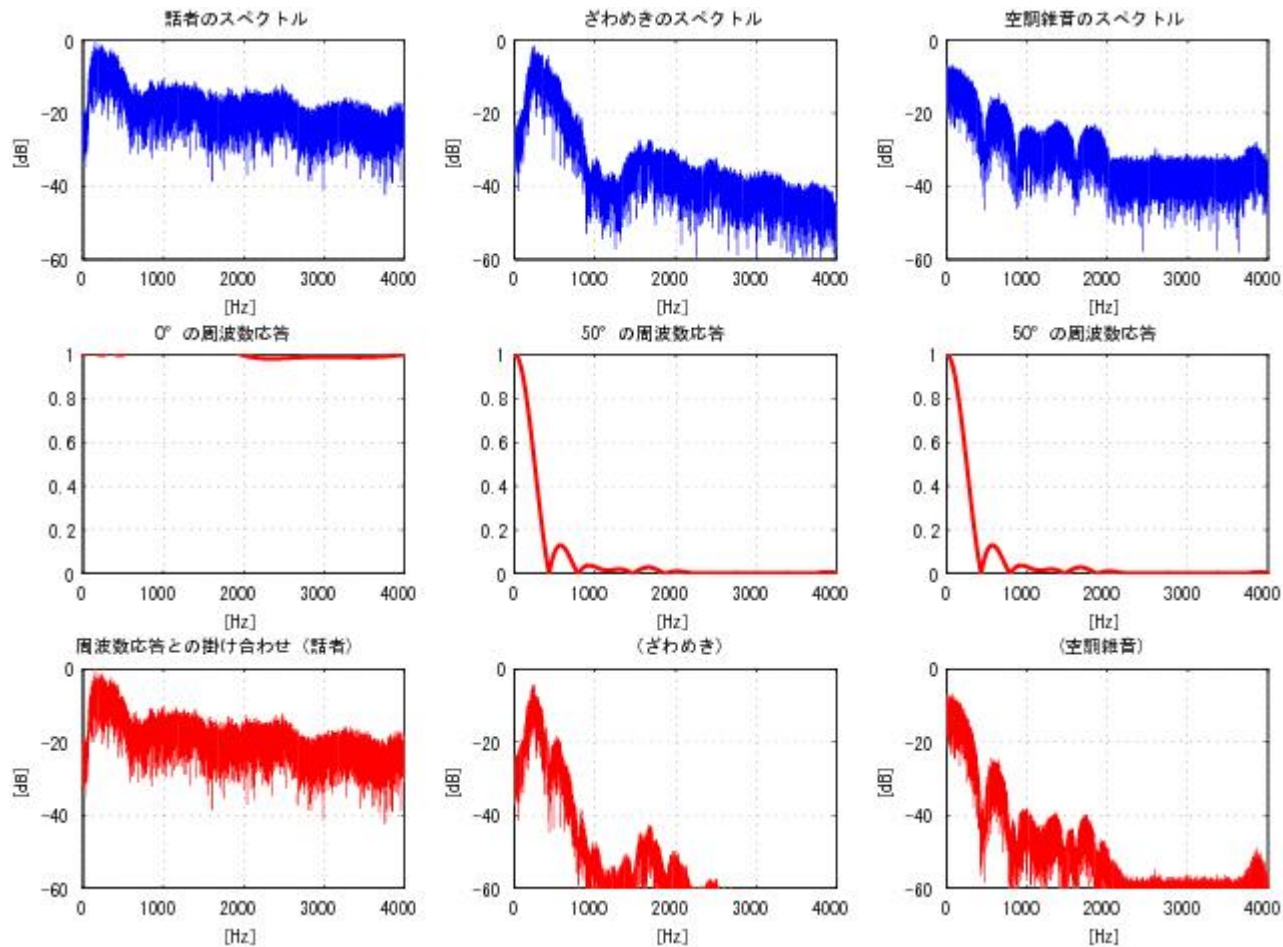
3. シミュレーション

3.3 シミュレーション結果



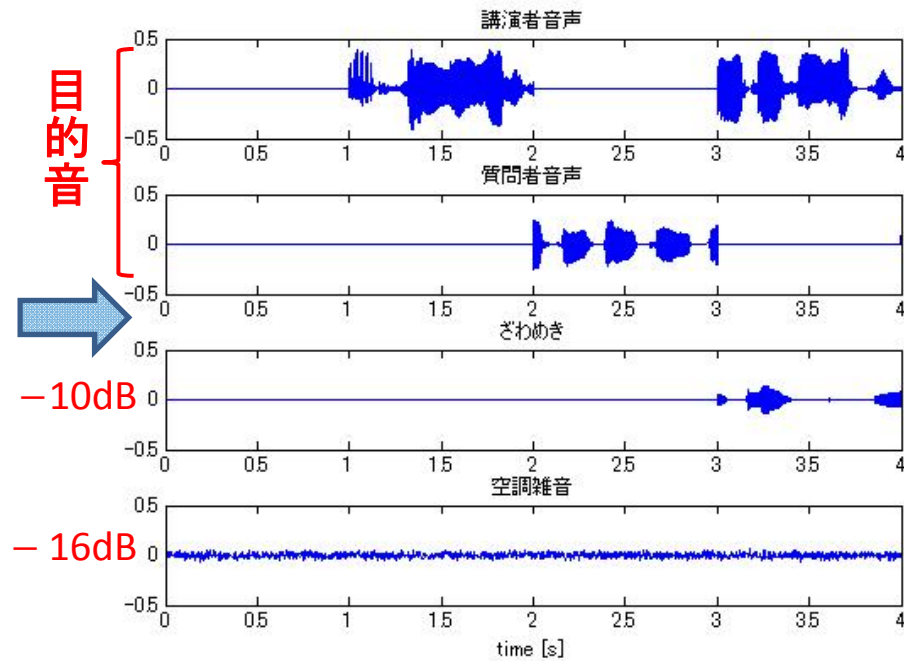
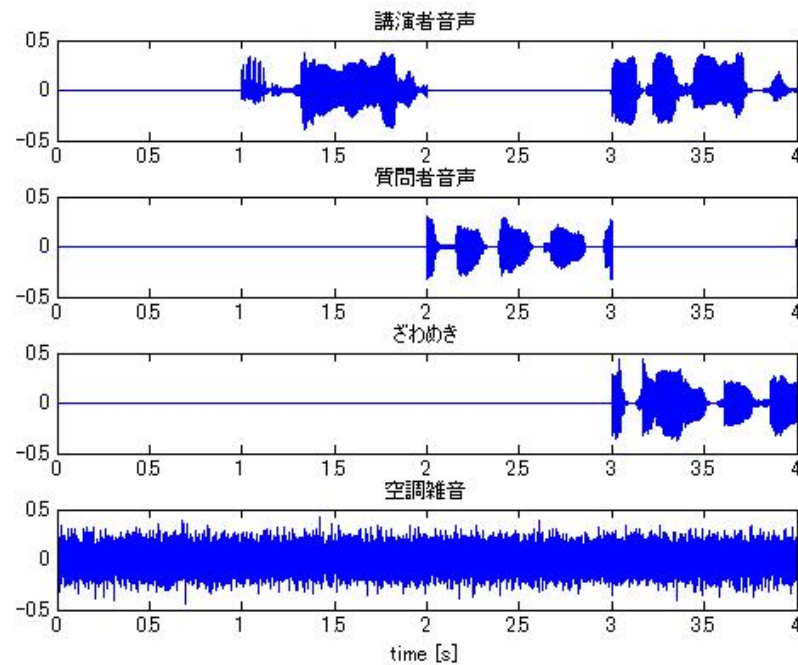
3. シミュレーション

3.3 シミュレーション結果



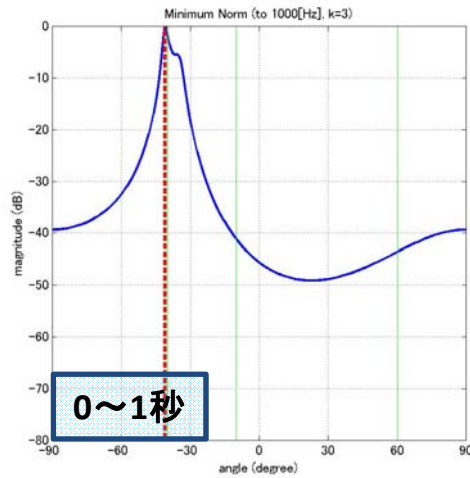
3. シミュレーション

3.3 シミュレーション結果



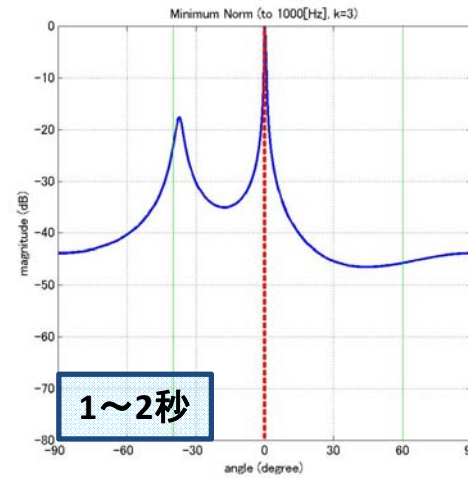
3. シミュレーション

3.3 シミュレーション結果



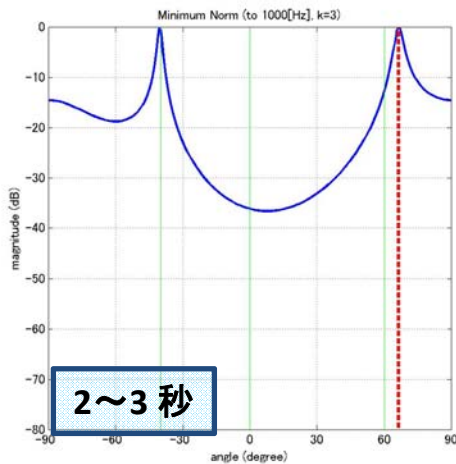
ESPRIT

-41.7°
0.9



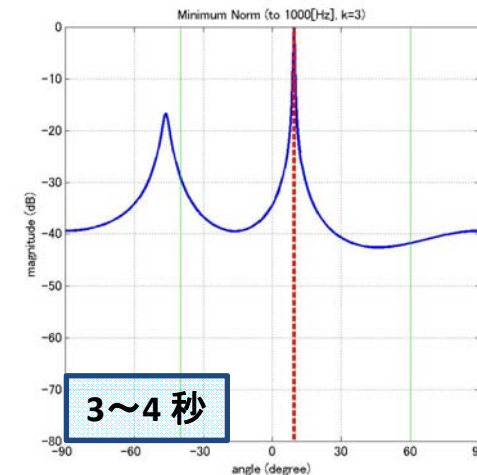
ESPRIT

0.1°
0.9
 -37.0°
0.1



ESPRIT

66.4°
0.7
 -40.5°
0.1



ESPRIT

16.1°
1.1
 35.1°
0.2
 -38.7°
0.1

3. シミュレーション

3.3 シミュレーション結果

< 各周波数での方向推定平均結果 >
(500, 600, 800, 1000, 1200, 1600, 2000, 2500, 3000[Hz])

(各時刻で最もパワーの大きい音源)

空調: -40° 講演者: 0° 質問者: 60° 講演者: 10°

時間[s]	0~1	1~2	2~3	3~4
最小ノルム	-40.6°	0.2°	64.5°	10.2°
ESPRIT	-40.6°	0.1°	64.3°	-0.8°

音源数がアレーの自由度を越えるときでも、
最小ノルム法では講演者音声の方向を高精度に推定できる

4.むすび

- 講演者音声のみを高S/N收音する
 - 指向性收音ビーム + 音源到来方向推定法

{シミュレーション実験}

1. 水平面の指向性形成
 2. 講演者音声のパワーを用いる方向推定
 3. 指向性收音ビームを講演者方向へ
- 今後の課題
 - 上下方向アレーの小型化
 - より実際的なシミュレーション

スピーカ系について

- 集束ビームを用いて生成した仮想音源の評価 -

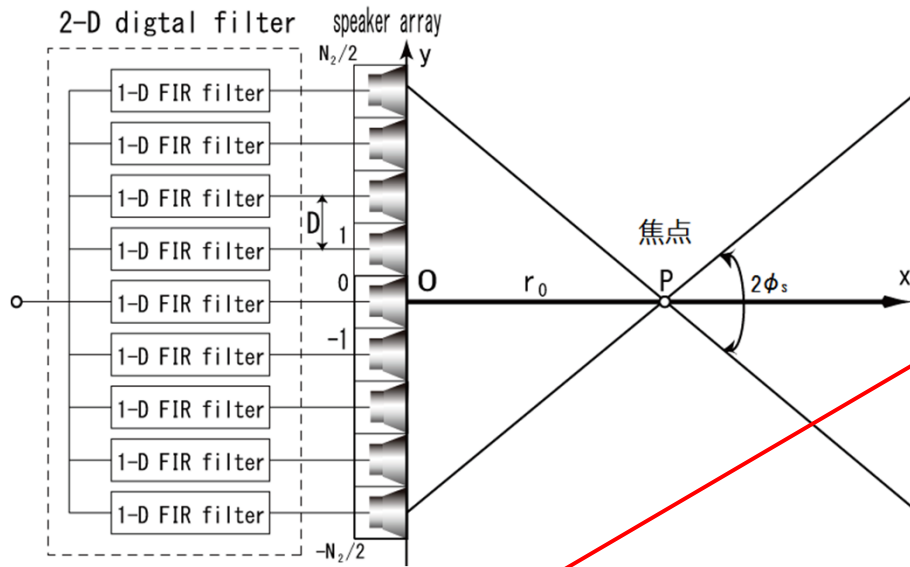
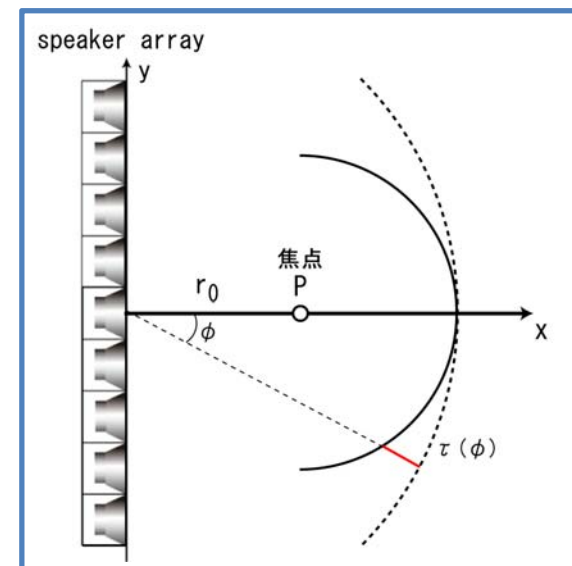
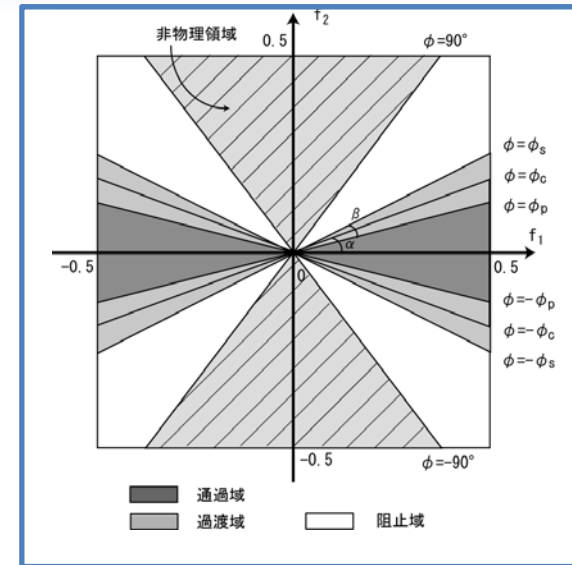


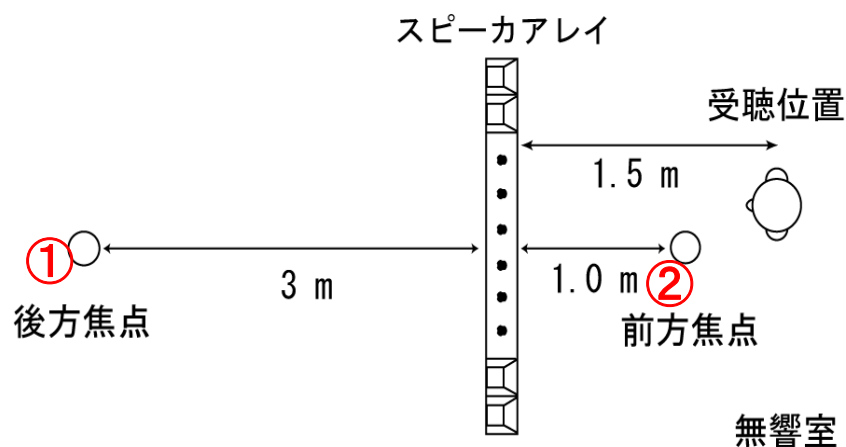
図 高音圧焦点の生成方法

振幅特性と位相特性の
制御により高音圧焦点を生成



音像制御

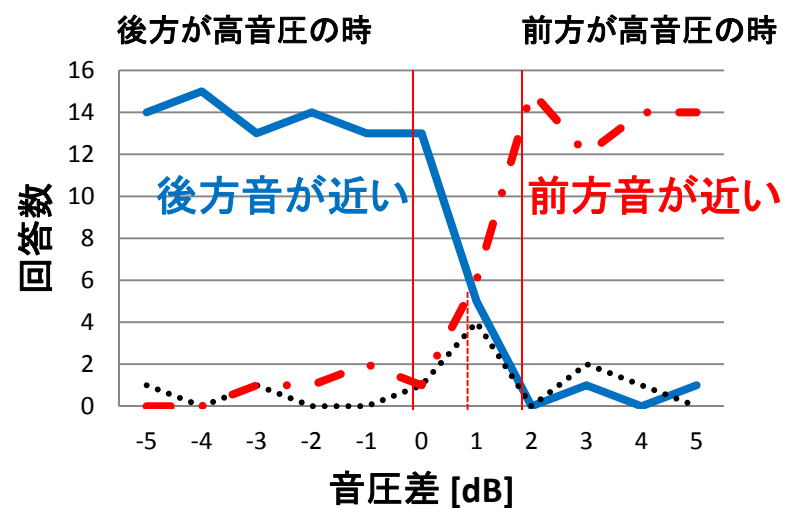
• 音圧差と距離感



聴取実験

- ・後方焦点から音を鳴らした後、前方焦点から音を鳴らす
- ・後方と前方の音圧差を変えていき、聴取者に音の遠近感を判断させる

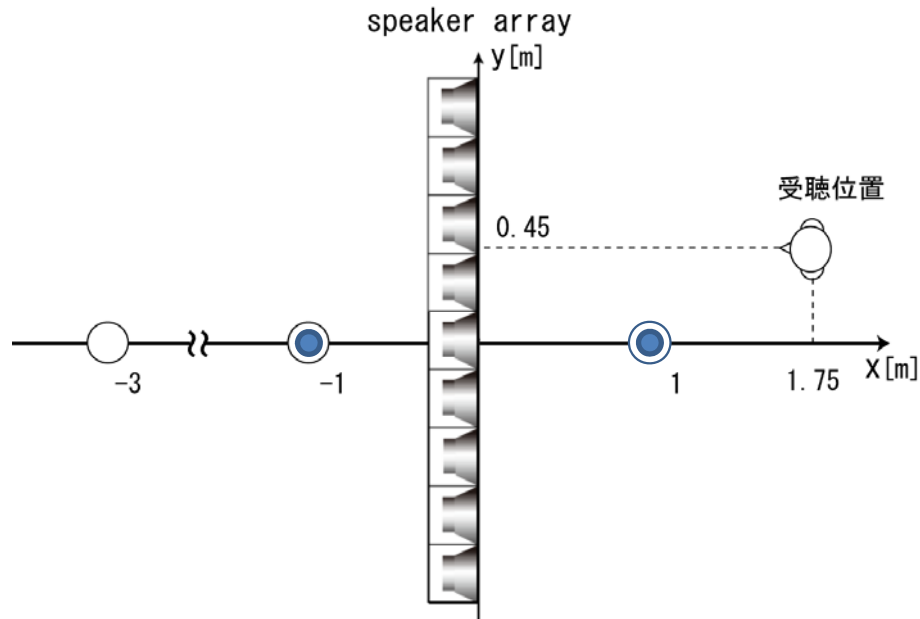
<回答結果>



遠近感 は 音圧差 1~2dB に依存

音像制御

● 波面と方向感・距離感

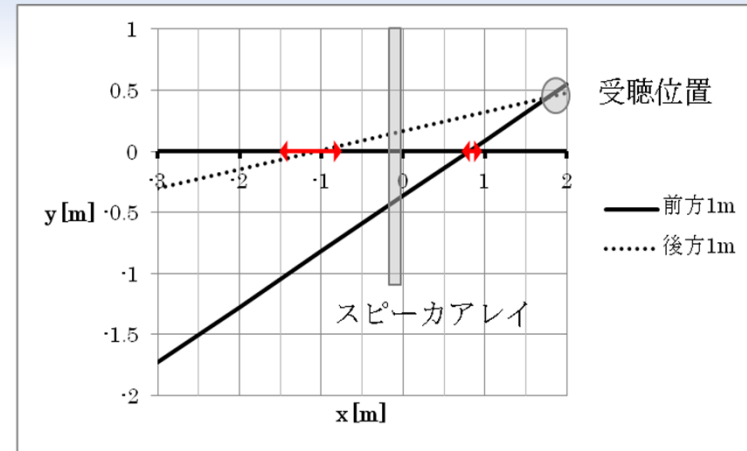


フィルタ次数 = (60, 14)

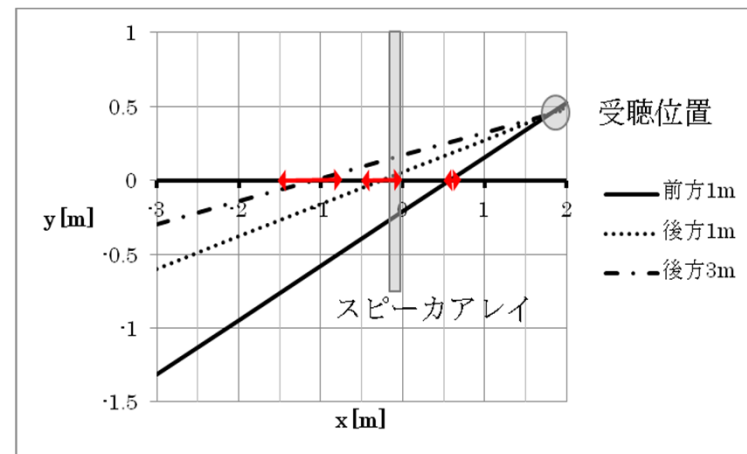
スピーカ間隔 = 14cm (スピーカ長 : 約2m)

信号帯域 = 2kHz

試験音 : 男声朗読音声「かえでいろづく」



実音源



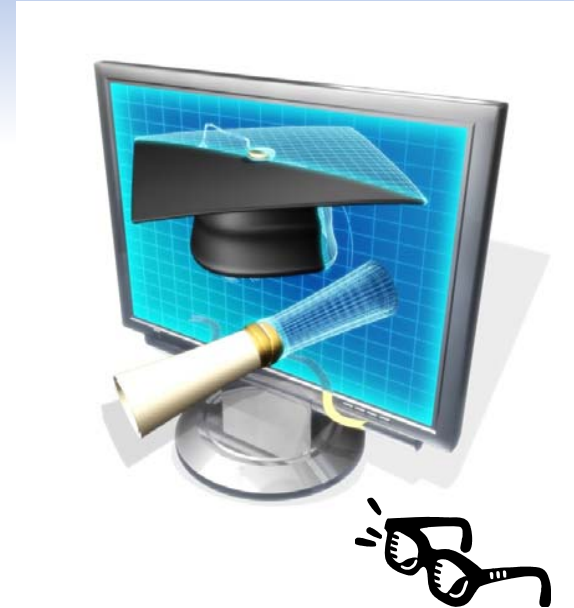
仮想音源

まえがき

近年、3D映像技術が発達している



奥行きはきわめてリアルに表現されている



音響の遠近感に関する研究では、

多数のスピーカを用いる**波面合成**や**焦点形成**が検討されている

我々は、

比較的小規模のスピーカアレイを用いた
焦点生成による音像制御方法の検討

まえがき

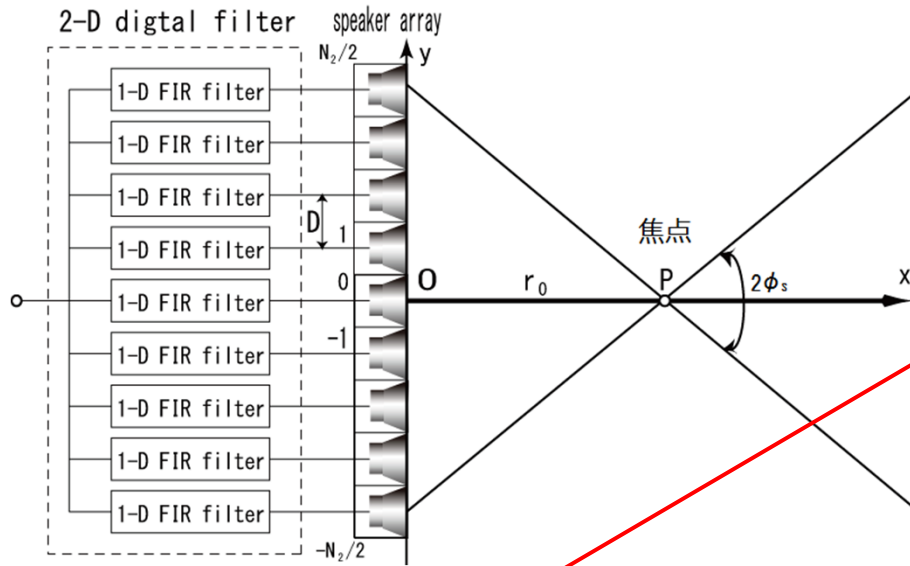
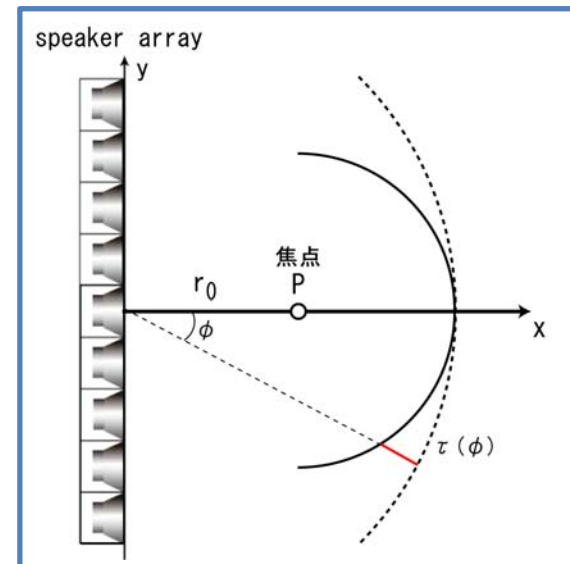
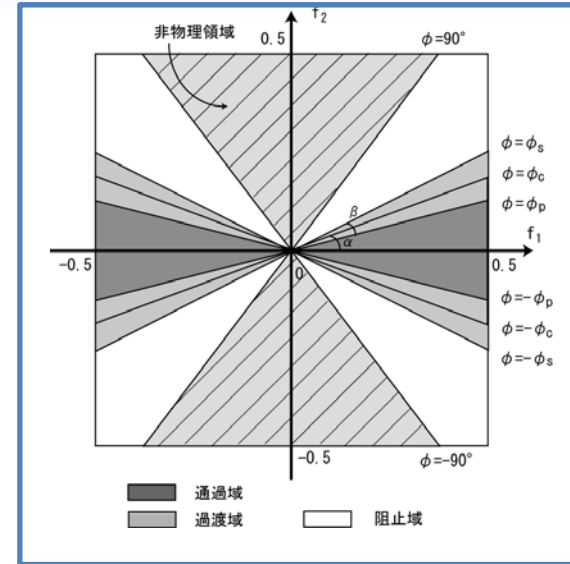


図 高音圧焦点の生成方法

振幅特性と位相特性の
制御により高音圧焦点を生成



まえがき

- 音源位置が異なった場合

- 音の大きさ



- 受聴位置での音圧差

- 波面の違い



- 受聴位置付近の位相差

- 壁による反射



- 直接音残響音比

スピーカアレイの前・後方に生成した仮想音源に関する聴取実験を行い、遠近感を評価する

遠近感について

- **音圧差**

遠近感に最も影響を与え、1~2 dB程度の差があれば十分影響を与える

- **波面**

実音源...低周波数の波面が遠近感に影響を与える

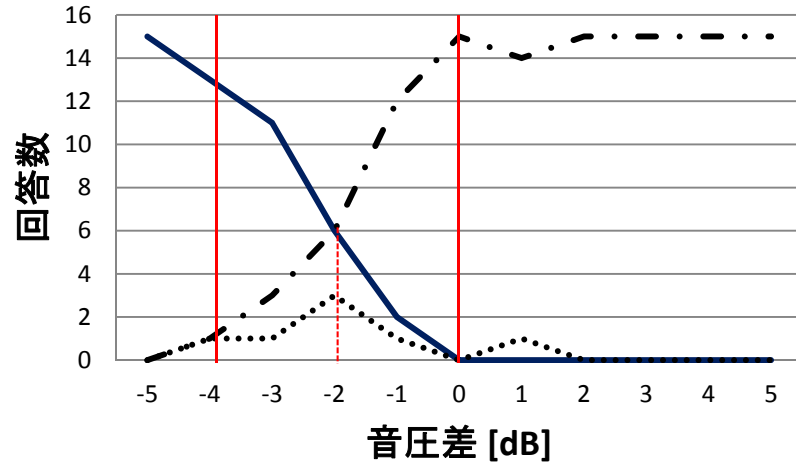
集束ビーム...低周波数の波面が形成出来ていないため検討は不十分

- **壁による反射**

遠近感に影響を与えるが、音圧差ほどの効果はない

音圧差について

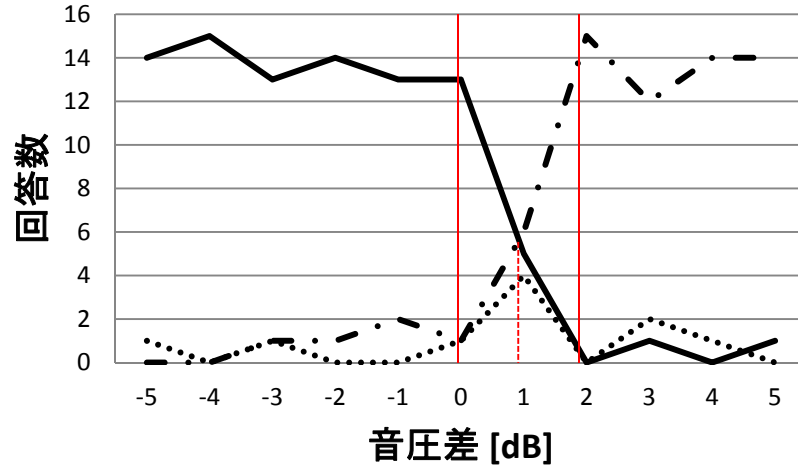
前方焦点音→後方焦点音



— 後続音の方が遠い
..... 遠近判別出来ない
- · - 後続音の方が近い

遠近判定の境目: -2 dB
回答が偏るまでの音圧差の幅: 約2 dB

後方焦点音→前方焦点音



遠近判定の境目: 1 dB
回答が偏るまでの音圧差の幅: 1 dB

↳ 遠近判断は
音圧差に依存
(1~2 dB)

遠近感について

- 音圧差

遠近感に最も影響を与え、1~2 dB程度の差があれば十分影響を与える

- 波面

実音源...低周波数の波面が遠近感に影響を与える

集束ビーム...低周波数の波面が形成出来ていないため検討は不十分

- 壁による反射

遠近感に影響を与えるが、音圧差ほどの効果はない

波面について

聴取実験を行った結果

実音源 ⇒ ○: **正確な遠近感**が得られる
集束ビーム ⇒ ×: 遠近感は得られなかった

先行研究より... **低周波数の波面**が聴感に影響を与えやすい



低周波数の波面を調べると...

実音源 ⇒ ○: 波面が**形成されている**
集束ビーム ⇒ ×: 狙いどおりの波面が**形成されていない**

本研究についても

低周波数の波面が遠近感に影響を与えることが示された

➡ 集束ビームについて波面を形成する必要がある

※ **アレイ長・スピーカ間隔**を検討


まとめ・今後の展望

- 集束ビームにより生成した仮想音源の遠近感に関する評価を行った
 - 遠近感は受聴位置での音圧差が最も影響を与える(1~2 dB)

不足点 : 低周波数における焦点の形成精度

今後は,

- 特に低周波数において高精度に波面を形成し, 波面の効果について再検討
 - アレイ長・スピーカ間隔を慎重に検討



ご清聴ありがとうございました