

包囲型マイクロフォンアレイを用いた音源位置推定に関する検討

岡本 拓磨, 西村 竜一, 岩谷 幸雄

東北大学 大学院情報科学研究科 / 電気通信研究所

はじめに

空間情報を含む音響技術の発達



よりリアルな音環境を体験できる

バーチャルヘッド

ある部屋に任意の仮想的な頭を想定し，そこで聞こえる音を再現する



このシステムの実現を目指す

具体的なシステムの流れ

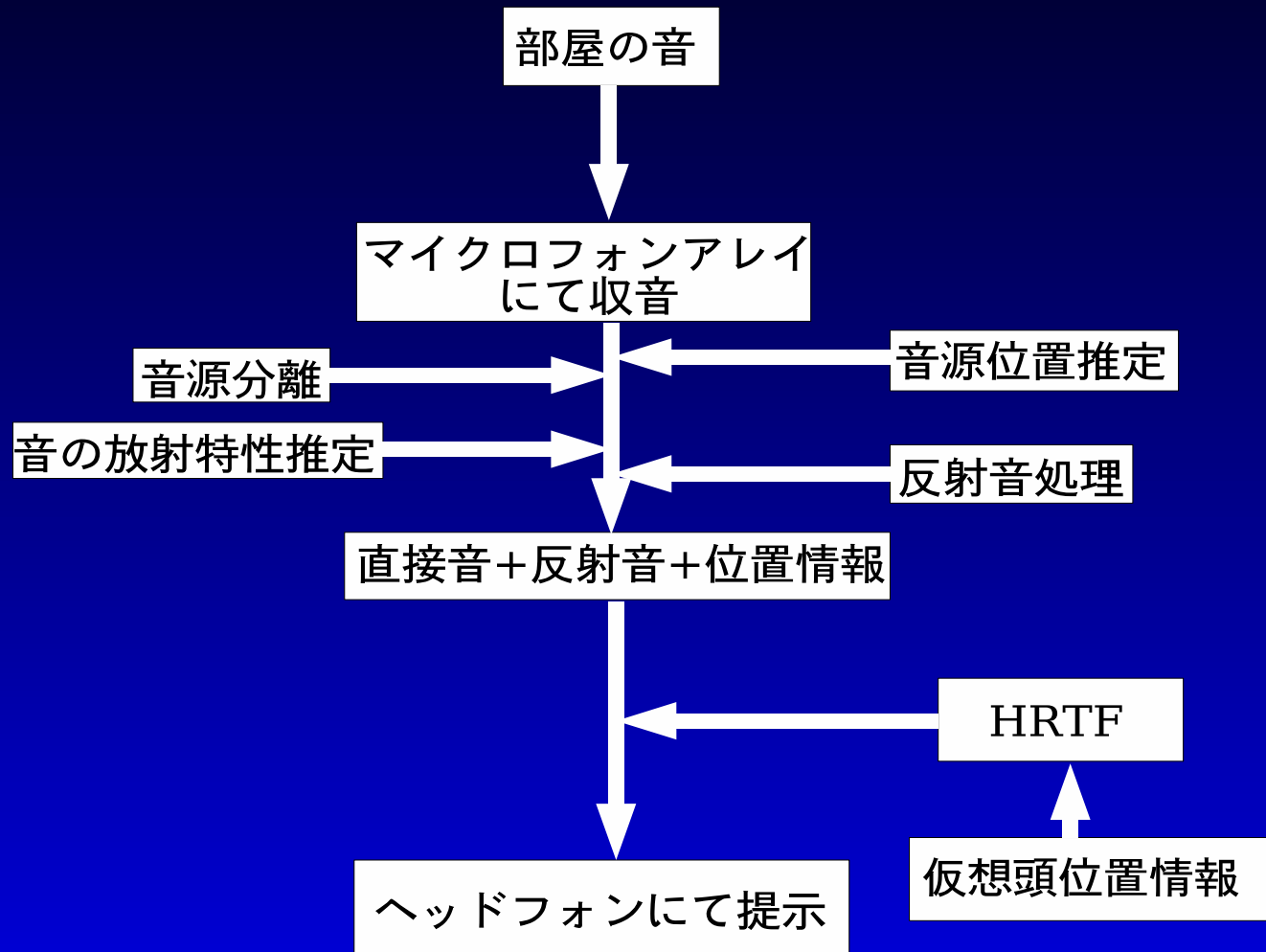


図 1 Virtual Head System

今回の目的

部屋の音をまるごと取りこむ



包囲型マイクロフォンアレイを使用

今回の目的

高精度な音源位置推定を行うための
マイクロフォン配置について検討

包囲型マイクロフォンアレイ

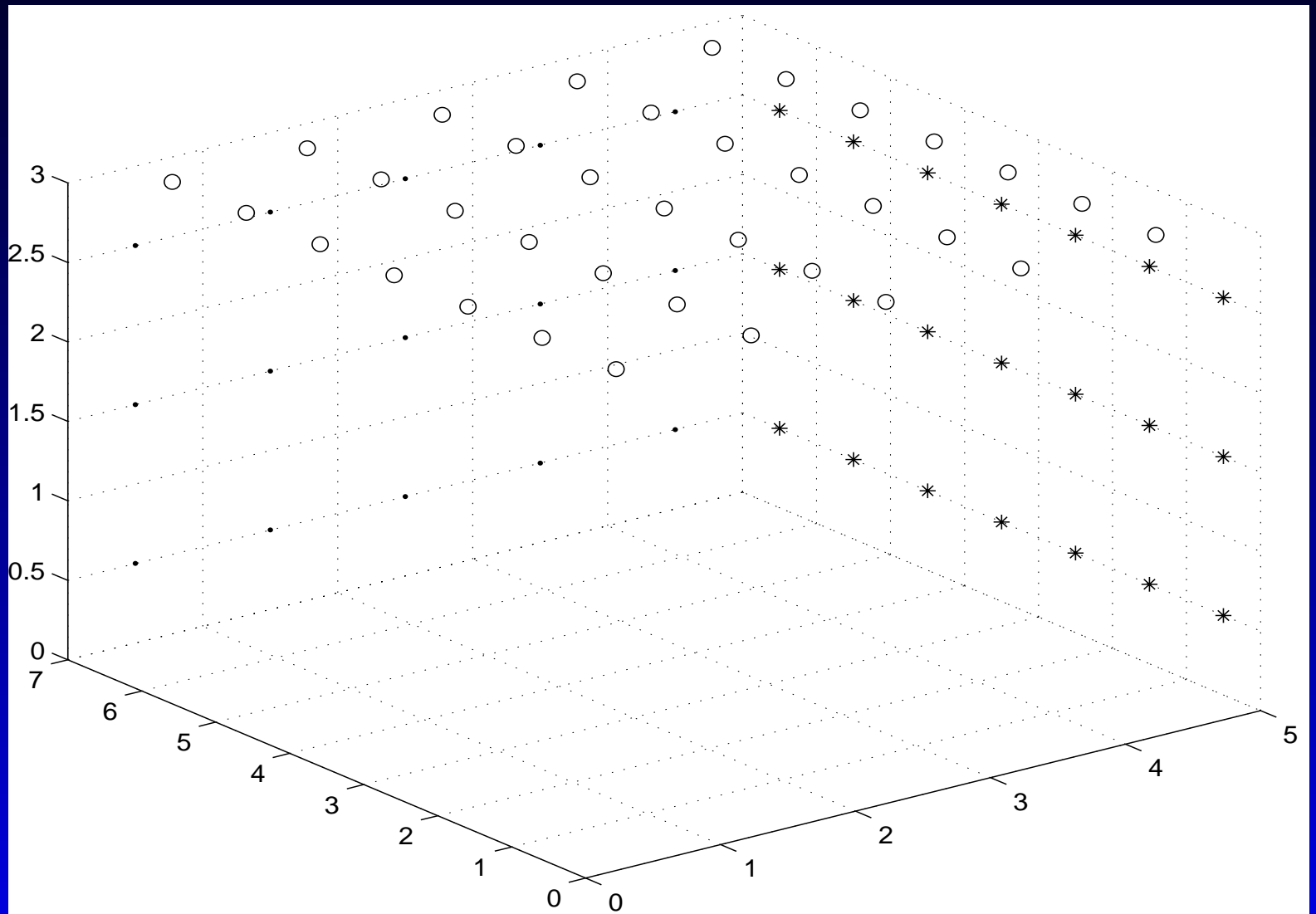


図2 包囲型マイクロフォンアレイ

マイクロフォンアレイの原理

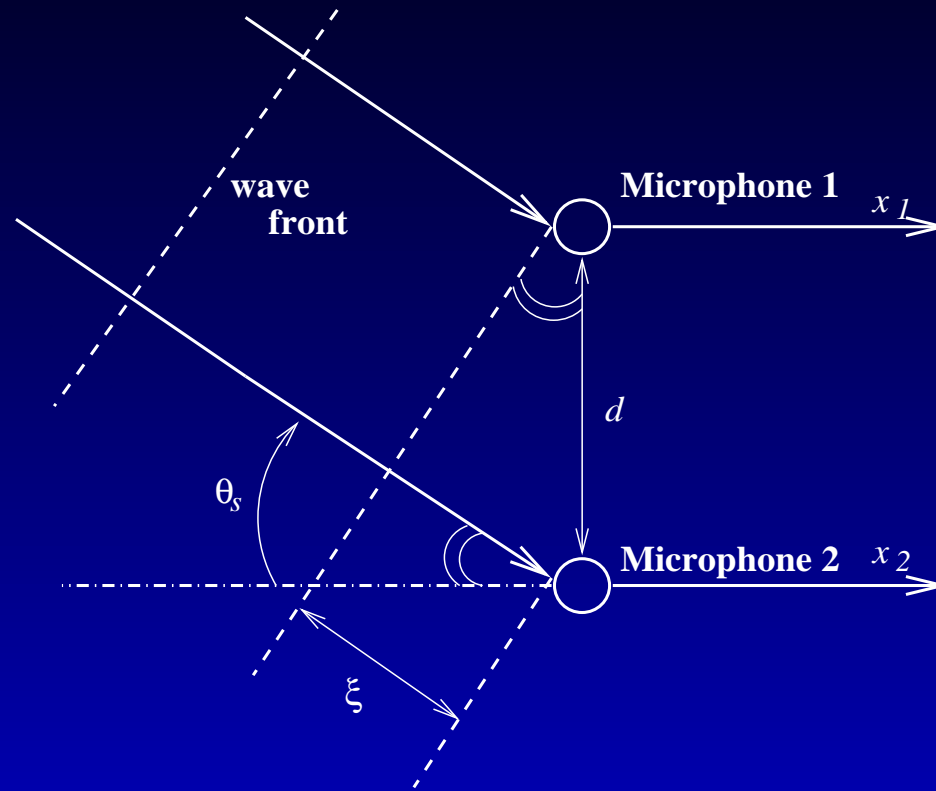


図 3 空間的に配置されたマイクロフォンで受音される信号

時間差 τ_s x_1 を τ_s だけ遅延 $x_1 + x_2$ θ_s 方向の信号を強調できる

遅延和アレイの原理

マイクログフォンアレイによる音源位置推定

- 遅延和を利用したものの
信号の到来時間差 τ_s を推定し，到来方向を推定する
- 最小分散法
注目している方向 θ_s 以外から到来する音の影響を最小にする
- Multiple Signal Classification (MUSIC) 法
到来方向には極めて高い分解能



今回は MUSIC 法を用いて音源位置推定を行う

MUSIC法(1)

入力ベクトル $x(\omega, t)$

$$(1) \quad x(\omega, t) = \mathbf{A}(\omega)\mathbf{s}(\omega, t) + \mathbf{n}(\omega, t)$$

空間相関行列 $\mathbf{R}(\omega)$

$$(2) \quad \mathbf{R}(\omega) = \mathbf{A}(\omega)\mathbf{P}(\omega)\mathbf{A}^H(\omega) + \sigma\mathbf{I}$$

相関行列 $\mathbf{R}(\omega)$ のランクは D , $\mathbf{R}(\omega)$ を固有値展開

$$(3) \quad \mathbf{R}(\omega) = \mathbf{V} \text{diag}[\lambda_1, \dots, \lambda_D, 0, \dots, 0] \mathbf{V}^{-1}$$

MUSIC法(2)

$VV^H = I$, $V^H = V^{-1}$ より, V により $R(\omega)$ を対角化

$$V^H R(\omega) V = [V^H A] P [V^H A]^H = \text{diag}(\lambda_1, \dots, \lambda_D, 0, \dots, 0)$$

(4)

(4)式より

$$(5) \quad [V_s^H A] P [V_s^H A]^H = \text{diag}(\lambda_1, \dots, \lambda_D)$$

$$(6) \quad [V_n^H A] P [V_n^H A]^H = \text{diag}(\sigma, \dots, \sigma)$$

↓
↓
 V と $A(\omega)$ が直交

MUSIC 法 (3)

$A(\omega)$ が張る信号部分空間 (signal subspace) に対して

- V_s は基底ベクトル
- V_n は直交補空間の基底ベクトル

MUSIC 法による空間スペクトル

$$(7) \quad P(\theta) = \frac{1}{\| \mathbf{V}_n^H \mathbf{a}(\theta) \|^2}$$

⇓
 $\mathbf{a}(\theta)$ が真の音源ベクトルと一致

⇓
 $P(\theta)$ は大きな値をもつ

音源位置推定シミュレーション

虚像法を用いて反射波を模擬
インパルス応答を生成



信号と畳み込み入力信号を得る



入力ベクトル $x(\omega, t)$ より, 空間相関行列 $R(\omega)$ を
求める



$R(\omega)$ と式 (4) より V を求め, V と $A(\omega)$ より $P(\omega)$
を計算



各周波数ごとに求めた $P(\omega)$ を平均して, それをそ
の位置における P とする.

シミュレーション条件

シミュレーション条件

サンプリング周波数 : 16000 Hz

部屋の広さ : 5(m) × 7(m) × 3(m)

音源1の座標 (2.42, 2.46, 1.50)

音源2の座標 (3.84, 4.86, 2.00) 共に白色雑音

反射係数 : 0.91

虚像法にて使用

虚像音源次数 n : 12

窓関数 : ハミング窓

窓長 : 512

使用する帯域 : 500~3000 Hz

大まかな音源位置推定

最初に部屋全体を大まかに走査して
音源のだいたいの位置を推定



周辺を精密に走査して音源位置を推定する



部屋全体を大まかに走査するための
マイクロフォン配置について検討

マイクロフォンモデル

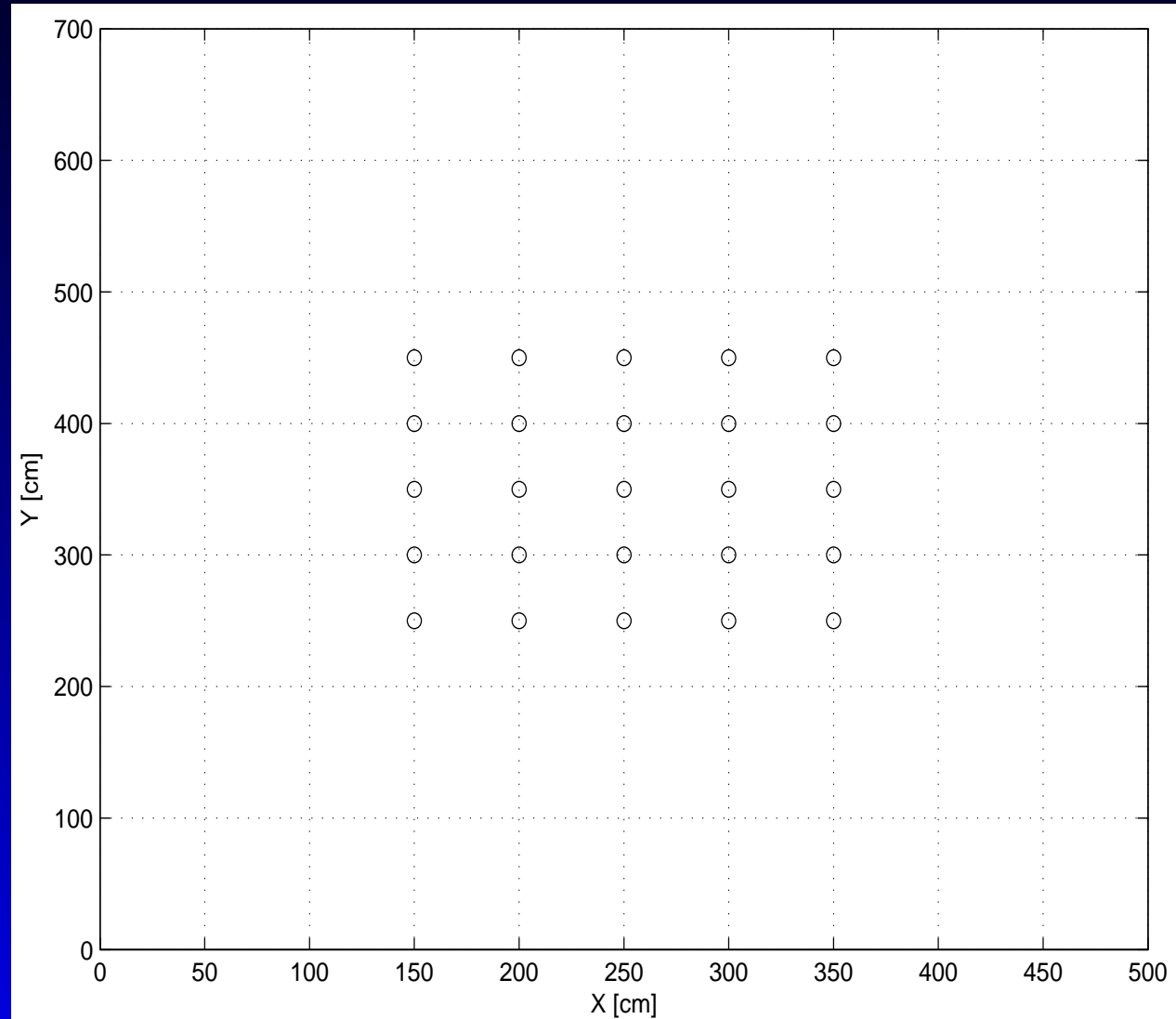


図4 パターン 0-1 (5 × 5=計 25 個 , 0.5(m) 間隔)

マイクロフォンモデル

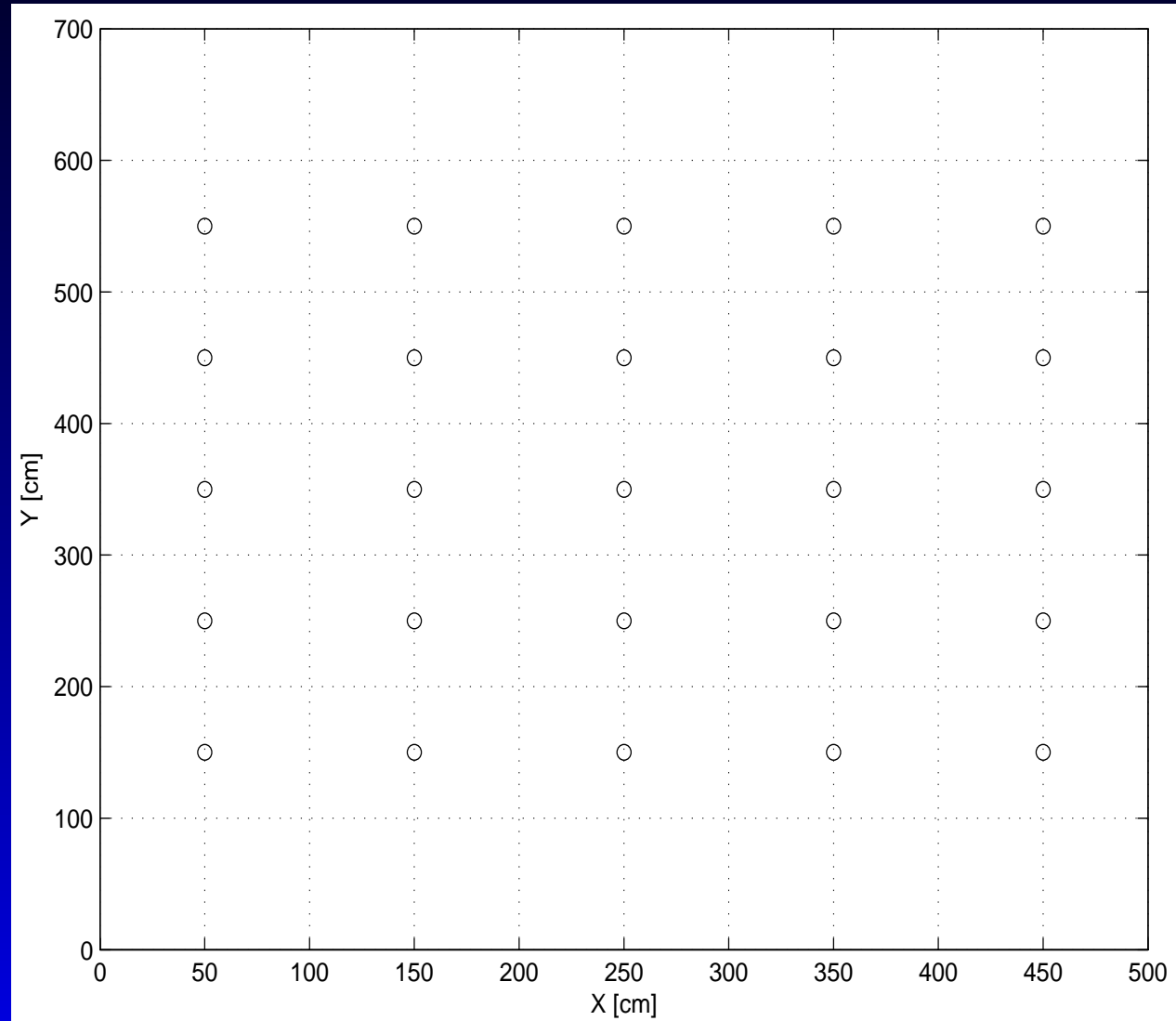


図5 パターン 0-2 (5 × 5=計 25 個 , 1(m) 間隔)

推定結果 (パターン 0-1)

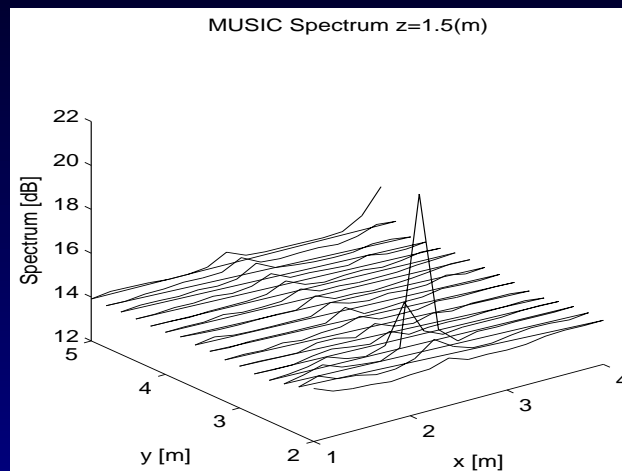


図6 パターン 0-1 結果 ($z=1.5$)

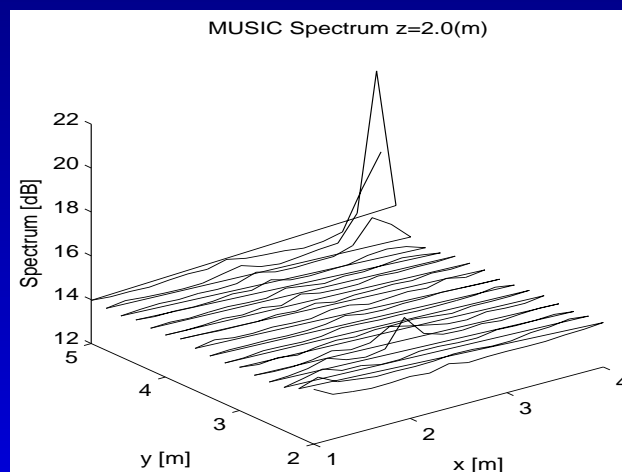


図7 パターン 0-1 結果 ($z=2.0$)

推定結果 (パターン 0-2)

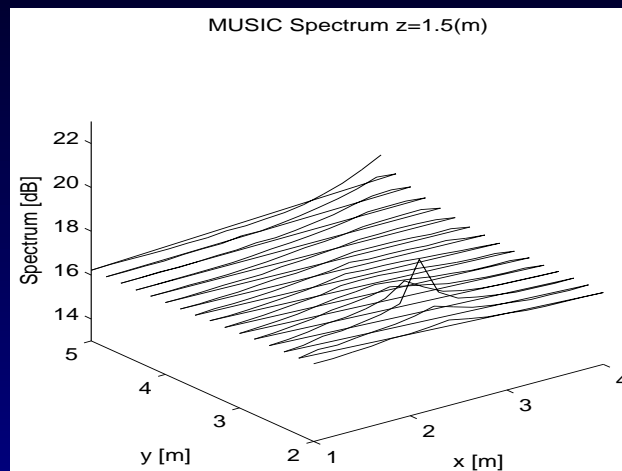


図 8 パターン 0-2 結果 ($z=1.5$)

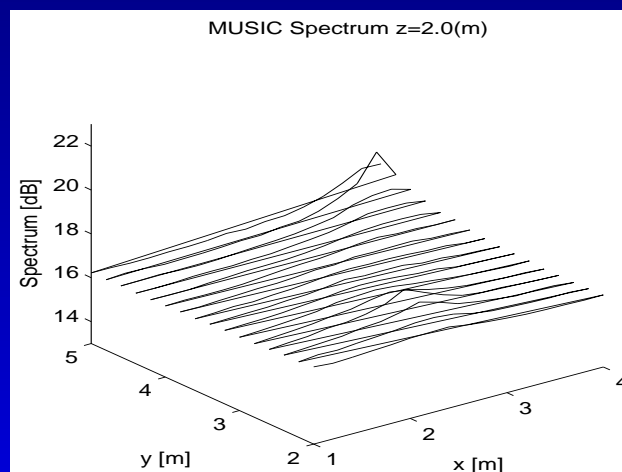


図 9 パターン 0-2 結果 ($z=2.0$)

詳細な音源位置推定

大まかな走査により得られたピークの周辺を
詳細に走査



3つの配置パターンで、 x, y, z それぞれ $\pm 0.4(\text{m})$,
5(cm) 間隔で
計 729 点を走査

パターン1

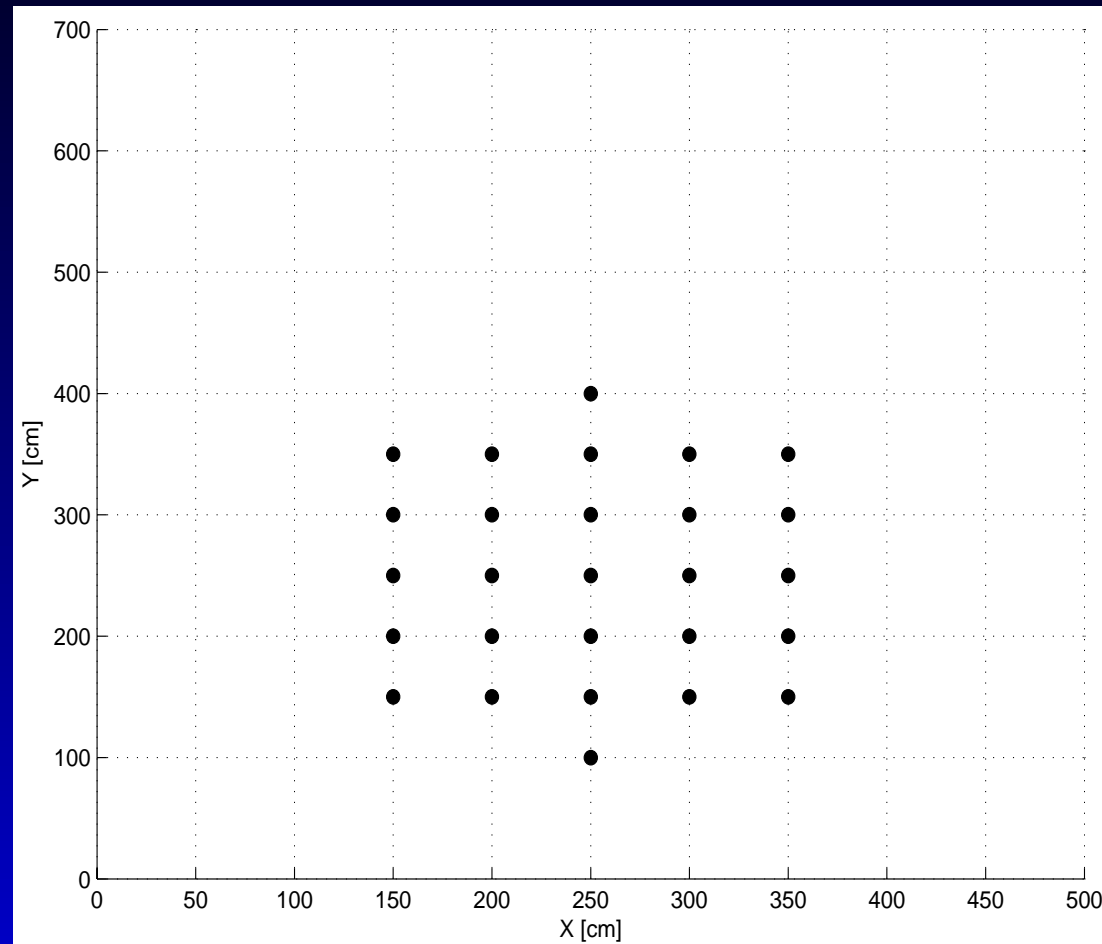


図 10 パターン 1

大まかに推定した x, y 座標がこのアレイのできるだけ中心にくるように天井に配置

パターン2

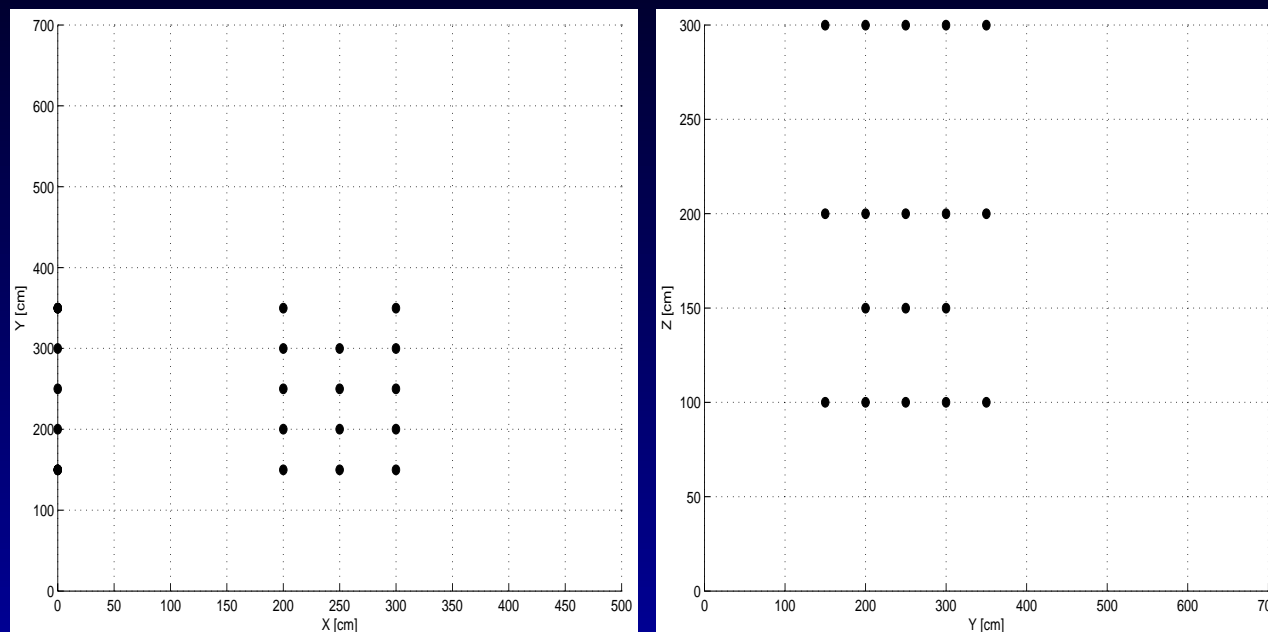


図 11 パターン 2

天井 ($z=3$) に 14 個 , 大まかに推定した x,y 座標のうち , もっとも近い壁に 13 個配置

パターン3

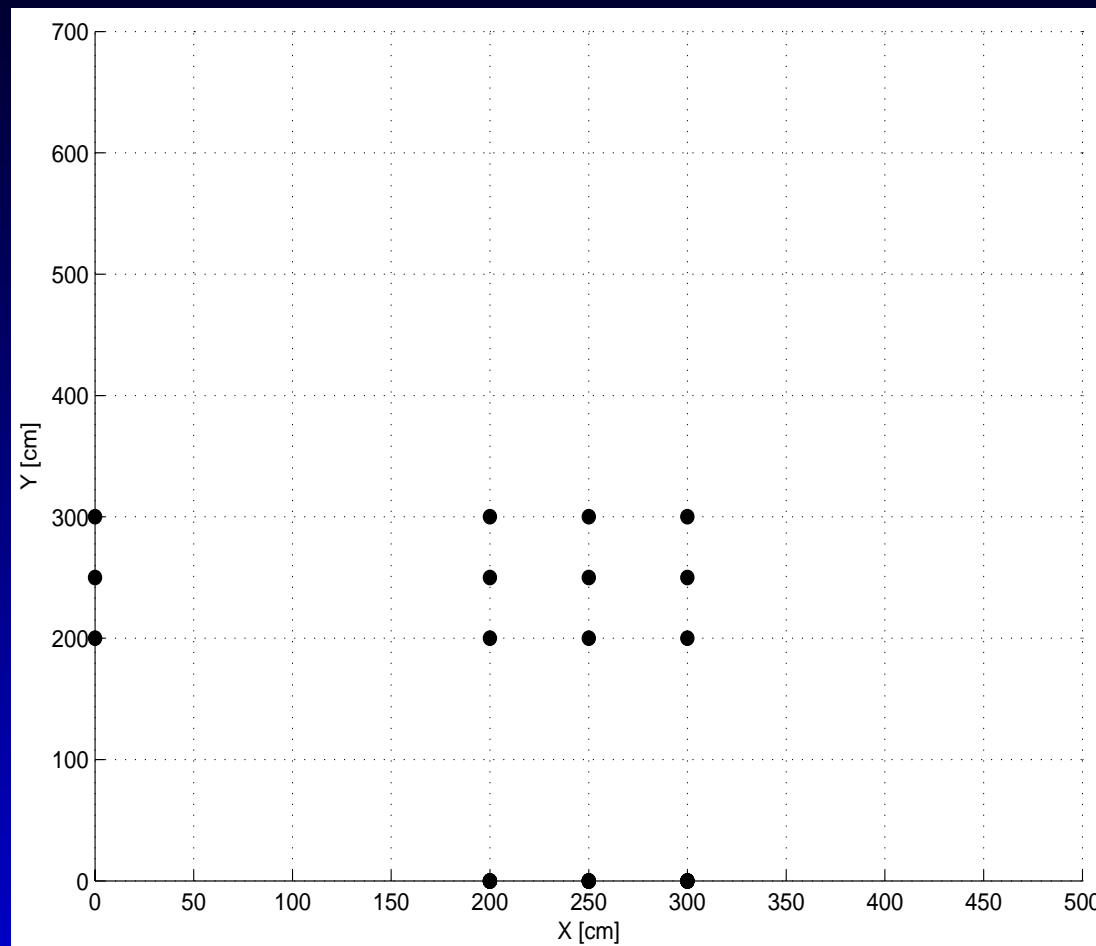


図 12 パターン 3 (z 軸方向から見た場合)
天井, 大まかに推定した x,y 座標それぞれに近い壁
に, 0.5(m) 間隔 $3 \times 3=9$ 個ずつ配置

探査結果

音源 1(2.42, 2.44, 1.5) 探査結果 (パターン 1)

x 座標	y 座標	z 座標	P
2.40	2.45	1.45	244.9
2.40	2.45	1.50	234.1
2.40	2.45	1.40	219.7
2.45	2.45	1.45	210.5
2.45	2.45	1.45	200.5

正しいピークが得られていない



残響なしで計算



それでも正しいピークは得られない

探査結果

音源 2(2.42, 2.44, 1.5) 探査結果 (パターン 2)

x 座標	y 座標	z 座標	P
2.40	2.45	1.50	139.5
2.45	2.45	1.45	134.5
2.40	2.40	1.50	81.7
2.45	2.40	1.45	81.6
2.50	2.45	1.40	74.1



正しいピークが得られている

探查結果

音源 3(2.42, 2.44, 1.5) 探查結果 (パターン 3)

x 座標	y 座標	z 座標	P
2.45	2.45	1.45	101.6
2.40	2.45	1.50	84.6
2.45	2.45	1.50	84.0
2.40	2.40	1.50	82.2
2.40	2.50	1.45	72.2

正しいピークが得られていない



残響なしで計算



それでも正しいピークは得られない

パターン3について

マイクロホン間隔を倍の1(m)にして再度推定
音源1 探査結果 (パターン3) アレイ間隔1(m)

x 座標	y 座標	z 座標	P
2.40	2.45	1.50	65.9
2.40	2.40	1.50	59.8
2.45	2.45	1.50	59.2
2.45	2.45	1.45	53.3
2.40	2.40	1.55	48.1



正しいピークが得られている

その他2条件

音源2についてはどれも正しい値を推定



以下の2つの条件についても推定

- 部屋の角に音源1を移動
音源3(1.5, 1.5, 1.5)
パターン3が最も得意としそうな場所なため
- 音源が部屋の下側 ($z < 1.5$) にある場合
音源4(2.42, 2.44, 1.0)
 z 軸だけは床 ($z=0$) にマイクロホンは設置できないため

音源3の推定結果

音源3(1.5, 1.5, 1.5) 探査結果 (パターン 1)

x 座標	y 座標	z 座標	P
1.50	1.50	1.45	133.8
1.50	1.50	1.40	124.6
1.50	1.50	1.50	122.0
1.50	1.50	1.35	106.1
1.50	1.50	1.55	99.9

音源3の推定結果

音源3(1.5, 1.5, 1.5) 探査結果 (パターン 2)

x 座標	y 座標	z 座標	P
1.50	1.50	1.50	154.3
1.55	1.50	1.45	102.4
1.45	1.50	1.55	87.3
1.50	1.55	1.50	83.2
1.55	1.55	1.45	73.2

音源3の推定結果

音源3(1.5, 1.5, 1.5) 探査結果 (パターン 3)

x 座標	y 座標	z 座標	P
1.50	1.50	1.50	223.2
1.55	1.55	1.45	97.3
1.45	1.45	1.55	82.9
1.50	1.55	1.50	63.3
1.50	1.50	1.45	63.1

音源4の推定結果

音源4(2.42, 2.44, 1.0) 探査結果 (パターン 1)

x 座標	y 座標	z 座標	P
2.40	2.45	0.95	146.4
2.40	2.45	1.00	143.7
2.40	2.45	0.90	141.8
2.45	2.45	1.00	136.9
2.45	2.45	0.95	134.6

音源4の推定結果

音源4(2.42, 2.44, 1.0) 探査結果 (パターン 2)

x 座標	y 座標	z 座標	P
2.40	2.45	1.00	96.5
2.45	2.45	0.95	90.3
2.40	2.40	1.00	69.5
2.35	2.45	1.05	69.3
2.45	2.40	0.95	69.0

音源4の推定結果

音源4(2.42, 2.44, 1.0) 探査結果 (パターン3)
アレイ間隔 1(m)

x 座標	y 座標	z 座標	P
2.40	2.45	1.00	66.0
2.45	2.45	0.95	58.0
2.45	2.45	1.00	53.5
2.40	2.40	1.00	50.6
2.40	2.40	1.05	49.6

結果と考察(1)

- パターン 1(天井 1 面)
 - 天井にしかマイクロフォンがないため, x, y 方向については精度よく推定できるが, z 方向の推定能力は低い
 - 音源が天井に近い場合は精度よく推定可能
- パターン 2(天井+壁 1 面)
 - 2面からの情報をクロスさせることにより, z 方向に関しても壁に配置されたアレイにより正しく推定されている

結果と考察(2)

- パターン3(天井+壁2面)
 - 音源が壁から離れている場合はマイクロフォン間隔を大きくしなければならない
 - 音源が壁に近いときはパターン2と同等の精度
 - 音源が部屋の角にあるときはもっとも良い精度
- 床にはアレイを設置できないことについて
 - パターン2とパターン3を用いれは音源が低い位置にあっても正しく推定できる

まとめ

今回のシミュレーションにおいて，高精度な音源位置推定を行うためには，壁側にもマイクロフォンアレイを設置することにより，天井だけに設置するよりも，特に部屋の高さ方向を高精度に推定できることが確認できた

今後の予定

- 今回のシミュレーション結果を実環境における実験により検討する
- 低周波を含む狭帯域音源に関する推定法を検討する
- 音源位置推定についても検討する



理論解についても検討していく