

包囲型マイクロホンアレイを用いた 音源位置推定および音源分離に関する検討

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

研究目標

ある部屋の音情報をまるごと取りこみ
そっくりそのまま別の場所で再現する

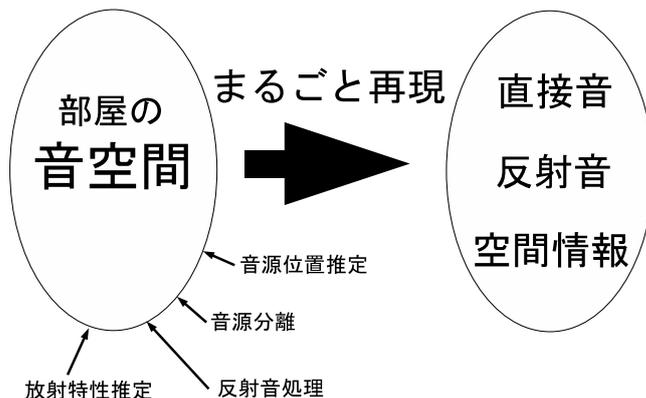


図 1: 時空間を超えた超現実的音空間の実現

システムの流れ

包囲型マイクロホンで部屋の音をまるごと取りこむ



- 音源位置推定
部屋のどこで音がなっているのか?
- 音源分離
どの種類の音がなっているのか?



より現実感を得るためには...

- 音の放射特性推定
人であれば顔の向きなど...
- 反射音処理
反射音の分離，呈示方法について...

包囲型マイクロホン

-床以外の天井、壁の計5面に
マイクロホンを格子状に設置

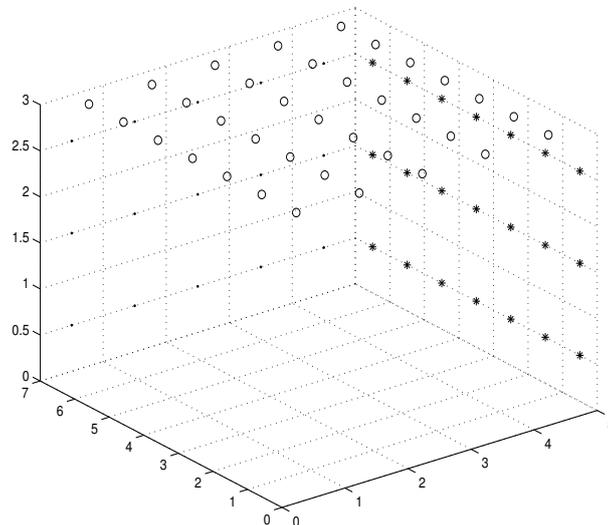


図2 包囲型マイクロホン (3面)

具体的には...

包囲型マイクロホンで部屋の收音



図 21 マイクロホン部屋@通研 W114



同じ配置のスピーカ部屋で再現



図 22 スピーカ部屋@通研 W114

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

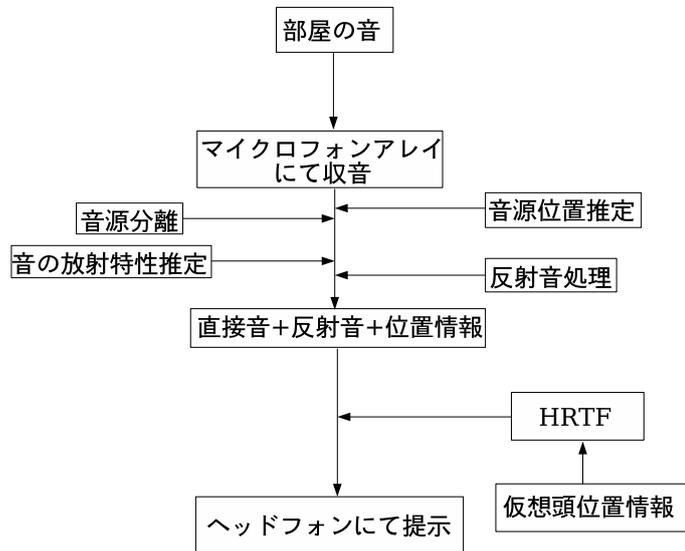


図 23 Virtual Head System

今回のテーマ

包囲型マイクロホンで部屋の
音をまるごと取りこむ



- **音源位置推定**

部屋のどこで音がなっているのか?

- **音源分離**

どの種類の音がなっているのか?



より現実感を得るためには...

- 音の放射特性推定

人であれば顔の向きなど...

- 反射音処理

反射音の分離, 呈示方法について...

包囲型マイクロホン

-床以外の天井、壁の計5面に
マイクロホンを格子状に設置

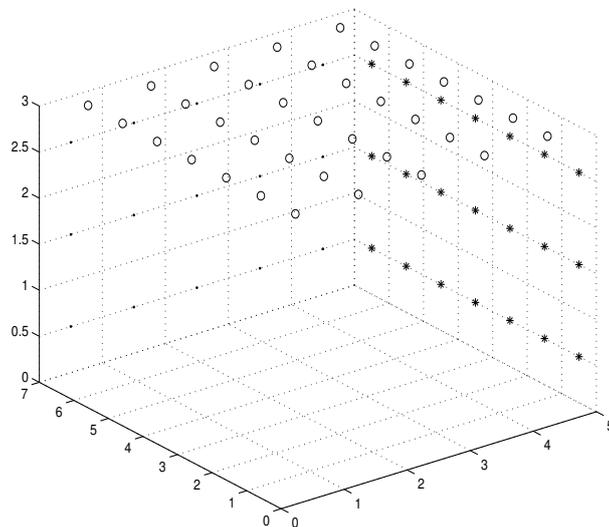


図2 包囲型マイクロホン (3面)

マイクロホンアレイの原理

各受信信号間には**到来時間差**や**振幅差**



これらの信号の差と**マイクロホンの位置情報**を利用



入力信号を処理することにより音源分離や音源位置推定が可能

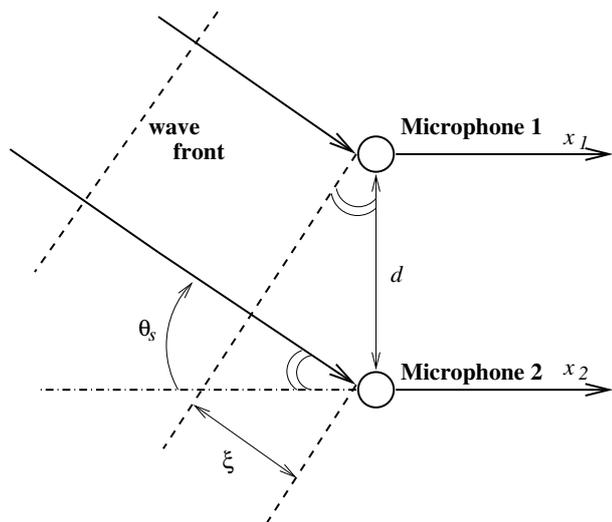


図3 遅延和アレイ

$$\text{時間差 } \tau_s = d \sin \theta_s$$



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

他方向からの信号は同相化されないため
強調されない

MUSIC法による音源位置推定

MUltiple SIgnal Classification(MUSIC)法

各マイクロホンの入力信号 $x(\omega, t)$ から, 各マイクロホンと音源位置の伝達関数 $A(\omega)$ と直交する直交補空間の基底ベクトル V_n を導出

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

から, θ 方向の伝達関数 $\mathbf{a}(\theta)$ でサーチして, $\mathbf{a}(\theta)$ が音源と同じ方向の時は V_n と直交するため $P(\theta)$ はピークをもつ



複数の音が存在する場合の分離機能や雑音耐性に優れている

マイクロホン配置が直線ではなく格子状など2次元的な場合, θ 方向の伝達関数 $\mathbf{a}(\theta)$ は $\mathbf{a}(x, y, z)$ となるため, 音源を3次元的にサーチすることが可能

包囲型マイクロホンによるMUSIC法シミュレーション

包囲型マイクロホンの特長 → 3次元的配置



MUSIC法の精度とマイクロホン配置について検討

- シミュレーションの流れ -

虚像法により得られた反射波と直接音から各マイクロホンの信号を生成



FFTした入力ベクトル $x(\omega, t)$ より, 直交補空間の基底ベクトル V_n を計算し, $a(x, y, z)$ との積を計算



各周波数ごとに求めた $P(\omega)$ を平均して, (x, y, z) における P とする

シミュレーション条件

- 部屋の広さ 3.0(m) × 5.0(m) × 2.1(m)
- 音源 1(音声)の座標 (1.20, 2.40, 1.00)
- 音源 2(音声)の座標 (1.60, 1.60, 1.50)
- サンプリング周波数 16000(Hz)
- 壁の反射係数 $\beta = 0.91$
- 虚像音源の次数 $n = 12$
- 入力信号の長さ 1 秒
- 窓関数 ハミング窓
- 窓長 512 点
- 周波数帯域 500 ~ 3000(Hz)

音声は NTT アドバンステクノロジー株式会社の親密度別単語理解度試験用音声データベースを使用

高精度推定のためのマイクロホン配置検討

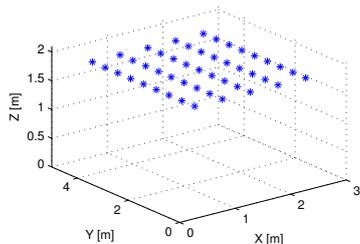


図4 パターン 1(1面)

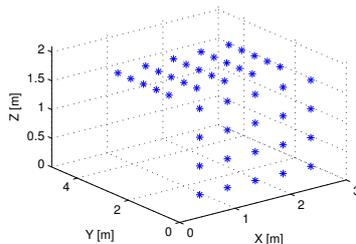


図5 パターン 2(2面)

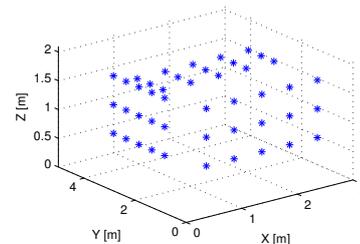


図6 パターン 3(3面)

表 1: 音源 1 の探査結果 (パターン 1)

x(m)	y(m)	z(m)	P	k
1.20	2.40	0.95	40.94	1.01
1.20	2.40	1.00	40.32	1
1.20	2.40	0.90	40.28	0.99
1.20	2.40	1.05	37.09	0.92
1.25	2.40	0.95	34.59	0.86

表 2: 音源 2 の探査結果 (パターン 1)

x(m)	y(m)	z(m)	P	k
1.60	1.60	1.50	38.48	1
1.60	1.60	1.45	37.04	0.96
1.60	1.60	1.55	35.83	0.93
1.60	1.60	1.40	33.23	0.86
1.60	1.60	1.60	32.18	0.85

表 3: 音源 1 の探査結果 (パターン 2)

x(m)	y(m)	z(m)	P	k
1.20	2.40	1.00	40.94	1
1.20	2.40	1.05	40.32	0.94
1.25	2.40	1.05	40.28	0.85
1.25	2.40	1.00	37.09	0.84
1.20	2.45	1.00	34.59	0.82

表 4: 音源 2 の探査結果 (パターン 2)

x(m)	y(m)	z(m)	P	k
1.60	1.60	1.50	40.05	1
1.60	1.60	1.45	34.17	0.85
1.60	1.60	1.55	33.81	0.84
1.65	1.60	1.50	29.67	0.74
1.55	1.60	1.50	29.36	0.73

表 5: 音源 1 の探査結果 (パターン 3)

x(m)	y(m)	z(m)	P	k
1.20	2.40	1.00	46.82	1
1.20	2.40	1.05	43.30	0.92
1.20	2.45	1.00	38.90	0.83
1.20	2.45	0.95	36.37	0.78
1.20	2.40	0.95	35.66	0.76

表 6: 音源 2 の探査結果 (パターン 3)

x(m)	y(m)	z(m)	P	k
1.60	1.60	1.50	40.71	1
1.60	1.60	1.45	36.09	0.89
1.60	1.60	1.55	35.68	0.88
1.60	1.60	1.60	29.67	0.73
1.65	1.60	1.50	29.63	0.73

結果と考察

- P が顕著なピークで現れない \Rightarrow 残響の影響
- パターン1 \Rightarrow 他と比べて推定能力が低い



2次元的な配置においてはその面と垂直な方向の推定能力が低い??

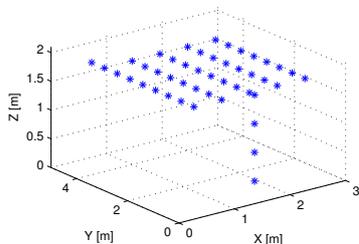


図7 パターン 4(1面と1直線)

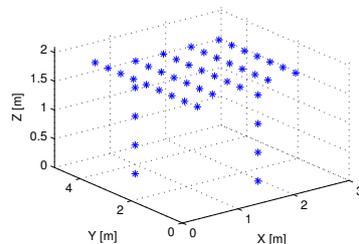


図8 パターン 5(1面と2直線)



パターン 2, 3 と同様な推定結果



マイクロホンを**3次元的に配置**することにより
2次元的に配置するよりも**高精度な推定結果**

シミュレーション結果の理論的考察

残響がない場合を考える

表 7: 音源 1 の探査結果 (パターン 1)

x(m)	y(m)	z(m)	P	k
1.20	2.40	1.00	831.27	1
1.20	2.40	1.05	521.64	0.62
1.20	2.40	0.95	340.27	0.41

表 9: 音源 1 の探査結果 (パターン 2)

x(m)	y(m)	z(m)	P	k
1.20	2.40	1.00	984.06	1
1.20	2.40	1.05	215.37	0.21
1.25	2.40	1.05	131.03	0.13

表 11: 音源 1 の探査結果 (パターン 3)

x(m)	y(m)	z(m)	P	k
1.20	2.40	1.00	933.64	1
1.20	2.40	1.05	221.01	0.23
1.20	2.40	0.95	131.27	0.14

表 8: 音源 2 の探査結果 (パターン 1)

x(m)	y(m)	z(m)	P	k
1.60	1.60	1.50	651.97	1
1.60	1.60	1.45	470.62	0.72
1.60	1.60	1.55	411.28	0.63

表 10: 音源 2 の探査結果 (パターン 2)

x(m)	y(m)	z(m)	P	k
1.60	1.60	1.50	957.47	1
1.60	1.60	1.45	168.78	0.17
1.60	1.60	1.55	143.64	0.15

表 12: 音源 2 の探査結果 (パターン 3)

x(m)	y(m)	z(m)	P	k
1.60	1.60	1.50	1058.2	1
1.60	1.60	1.55	232.29	0.21
1.60	1.60	1.45	218.83	0.21



どの結果も鋭いピークで推定



パターン 1(2次元配置)は他(3次元配置)に比べて
構成した面とは垂直方向の推定能力低い

伝達関数 $a(x, y, z)$ の類似性の検討

P がピークを形成しにくい



V_n はマイクロホン配置ごとに一定のため
音源位置の伝達関数 $a_s(x, y, z)$ と、その近くの点における
伝達関数 $a_s(x + dx, y + dy, z + dz)$ が類似している!!



伝達関数 $a_s(x, y, z)$ の類似性について検討

音源位置と各マイクロホンの伝達関数
 $a_s(x, y, z)$ と音源位置から z 方向に dz だけシフトした点と各マイクロホンの伝達関数
 $a_s(x, y, z + dz)$ の比較



それぞれ対応する距離ベクトル $d_m(x, y, z)$ 同士の大きさを比較

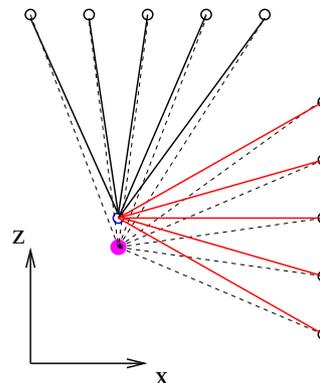


図 15 z 方向に 5(cm) シフト ベクトルの変化を比較

類似性の検討結果と考察(1)

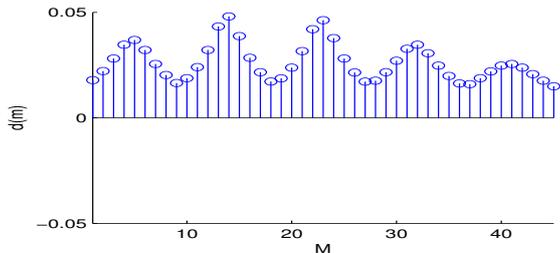


図9 パターン1でz方向に+5cmずらした場合

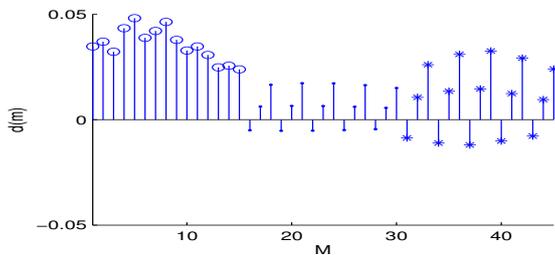


図11 パターン3でz方向に+5cmずらした場合

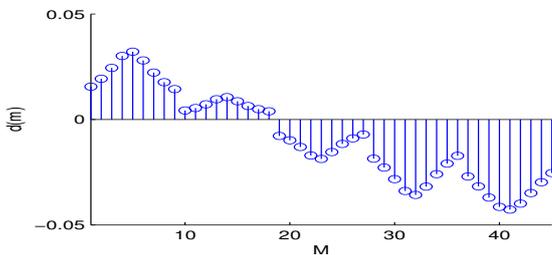


図13 パターン1でx方向に-5cmずらした場合

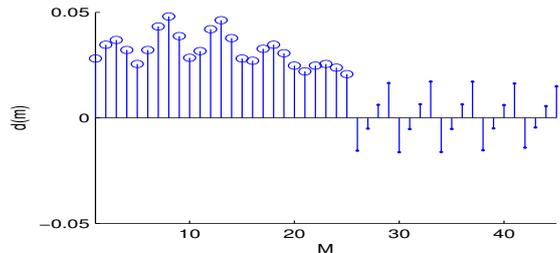


図10 パターン2でz方向に+5cmずらした場合

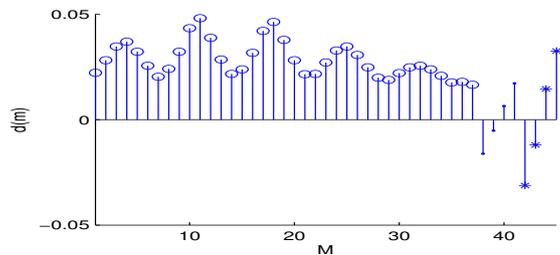


図12 パターン5でz方向に+5cmずらした場合

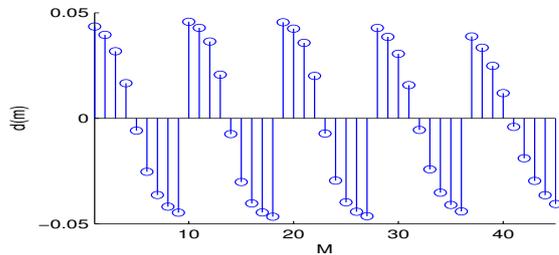


図14 パターン1でy方向に-5cmずらした場合

類似性の検討結果と考察(2)

パターン1のz方向にシフトした時だけ全て同じ符号向きにベクトルの大きさが変化する



パターン1の場合のz方向へのシフトが他に比べて一番数 $a_s(x, y, z)$ と近いものになる



これが原因でPのピークは出にくくなる \Rightarrow z方向の推定能力低い

その他のパターン \Rightarrow 天井以外のマイクロホンの存在により $a_s(x, y, z)$ には近くなりにくい



3次元配置の方が推定能力が高い

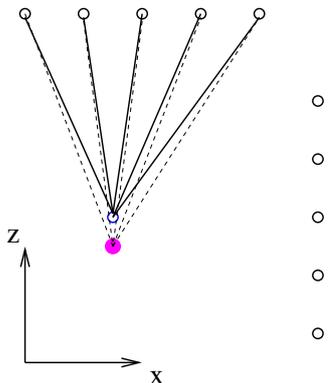


図 20 z 方向に 5(cm) シフト (パターン 1)

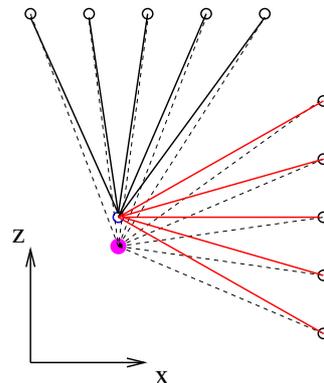


図 15 z 方向に 5(cm) シフト (パターン 2)

遅延和(3D)を用いた音源分離

- MUSIC法と同様, 3次元的なスポットを当てることが可能
- 音源から各マイクロホンへの伝搬時間を τ_m
- 固定遅延時間を D

各マイクロホンの入力信号を $(D - \tau_m)$ だけ遅延



音源方向からの信号を同相化



足し合わせると音源方向の信号のみを強調できる

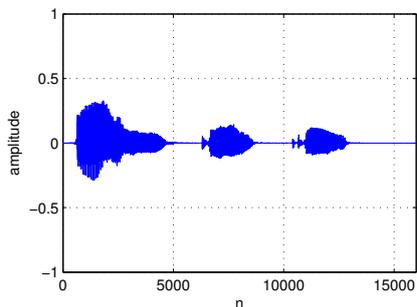


図 16 音源 1(fto) の原信号

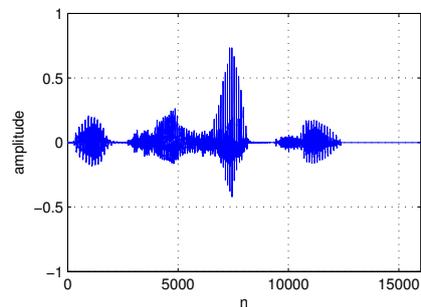


図 17 音源 2(mya) の原信号

結果と考察

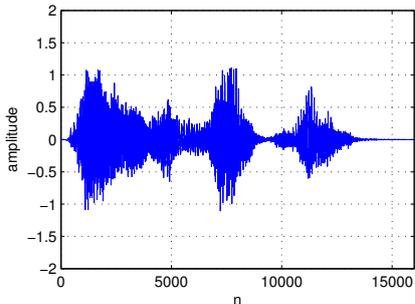


図 18 マイクロホン 1 における信号

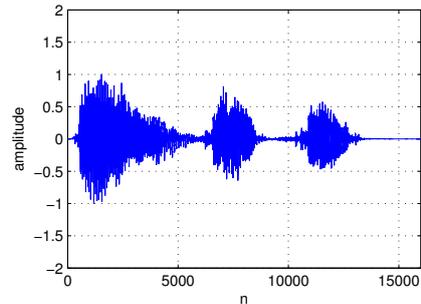


図 19 パターン 1 で音源 1(fto) を強調した場合

各パターンでの S/N 比の改善

パターン 1	音源 1 : 50.4dB	音源 2 : 28.0dB
パターン 2	音源 1 : 65.7dB	音源 2 : 24.4dB
パターン 3	音源 1 : 58.5dB	音源 2 : 32.3dB



S/N 比の改善率は面のとり方には依存していない



雑音の反射音がどのマイクにも音源と同じ方向から入っているため

まとめ

- ・シミュレーション結果より包囲型マイクロホンによる音源位置推定では天井の1面だけを用いるよりも，壁のマイクロホンも使用した2面，3面を用いた方が高精度であることを確認
- ・その理論的な背景を検討
- ・包囲型マイクロホンを用いた遅延処理による音源分離においては，面のとり方にはあまり依存しないことを確認

今後の予定

- ・実験のより今回の結果を確認
- ・実環境を想定した放射特性のある音についての位置推定，分離について検討