

EA研究会, 聴覚研究会

# 指向性スピーカと波面合成法を用いた 近接三次元音場再生システムの 性能比較評価

木村敏幸<sup>1</sup>, 山肩洋子<sup>2</sup>, 勝本道哲<sup>1</sup>,  
岡本拓磨<sup>3</sup>, 矢入聡<sup>4</sup>, 岩谷幸雄<sup>3</sup>, 鈴木陽一<sup>3</sup>

<sup>1</sup>NICT, <sup>2</sup>京都大学, <sup>3</sup>東北大学, <sup>4</sup>仙台高専

# 超臨場感コミュニケーション

- “未来の”3Dテレビ
  - リビングで鑑賞
  - 目の前に対象物がある
  - メガネがいらぬ
- 立体遠隔通信会議
  - 同じ場所で会議
  - 目の前に相手がいる
  - メガネがいらぬ
- 波面合成法に着目



リビングで楽しむ  
立体テレビ放送

立体映像による  
コミュニケーション

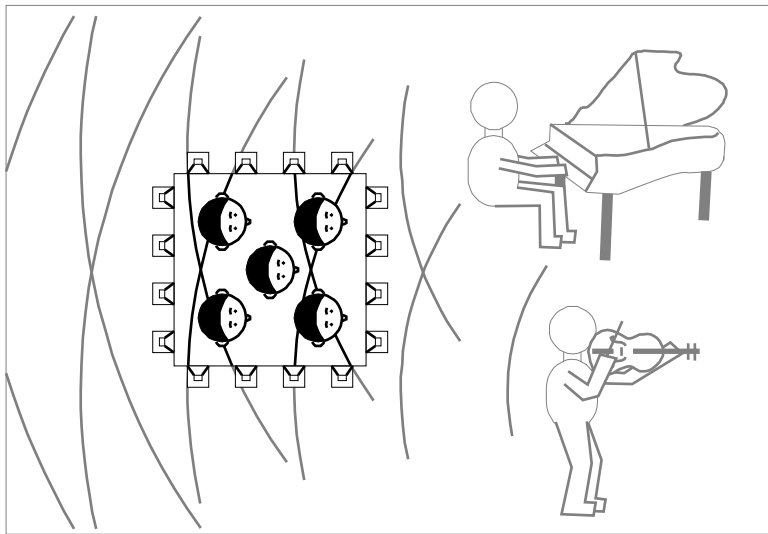


会場で新作サンプルの  
立体映像を遠隔地と共有

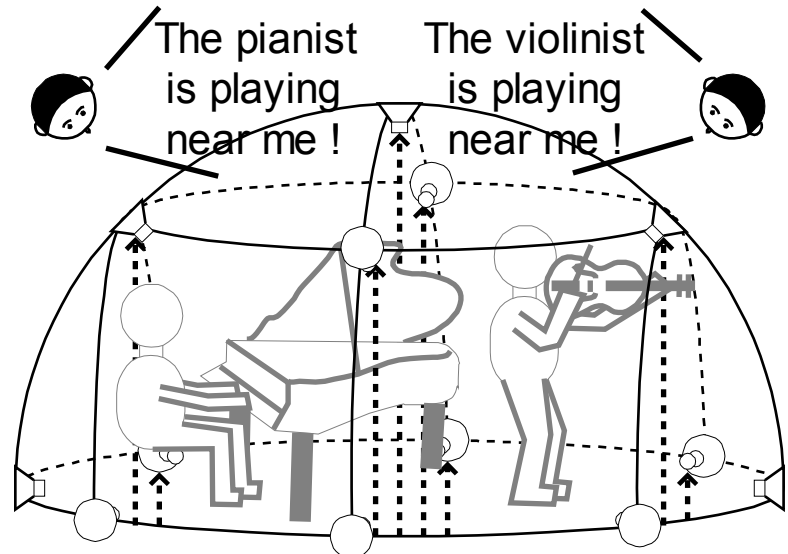
# “周囲から聴く”立体音響システム

- 今までの波面合成システム
  - 聴取者の周りにスピーカーを配置
- 提案する波面合成システム
  - 音源の周りにスピーカーを配置

Conventional System



Proposed System



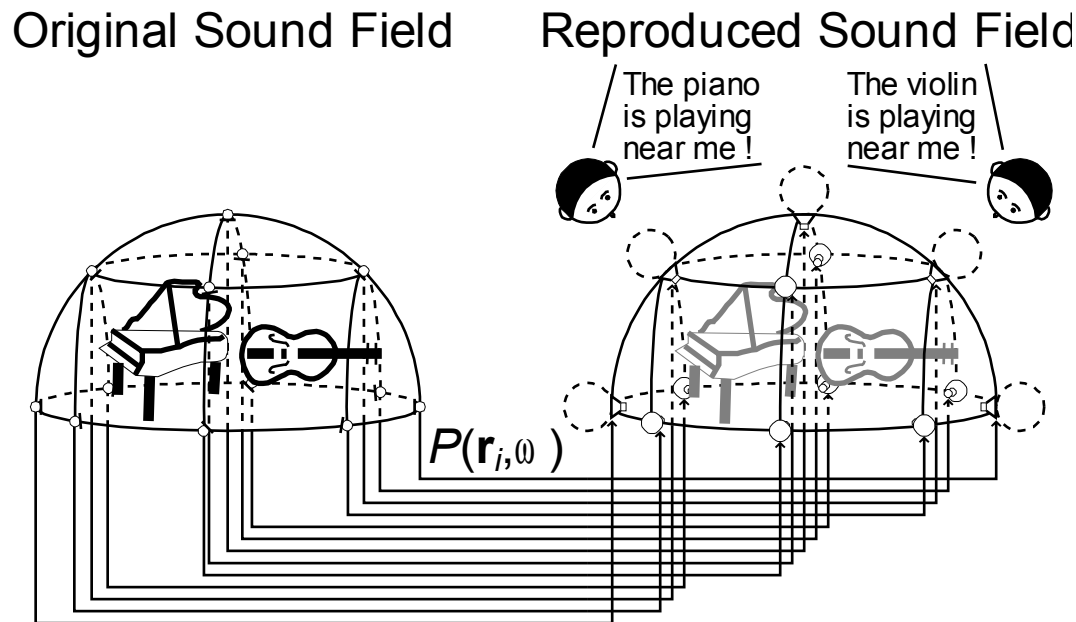
# 本報告の目的

---

- 指向性スピーカと波面合成法を用いた近接三次元音場再生システム
  - 今までに提案し, 実装
- 実装したシステムの音像定位性能の評価
  - 音像位置推定を実施し, 結果を比較
    - 計算機シミュレーション
    - 音響測定

# 提案システムの概要

1. 境界面上のマイクロホンで音を収録
2. 境界面上の指向性スピーカで音を再生
3. アレイの外側の音場が再現
4. アレイの内側で音が鳴っているように感じる



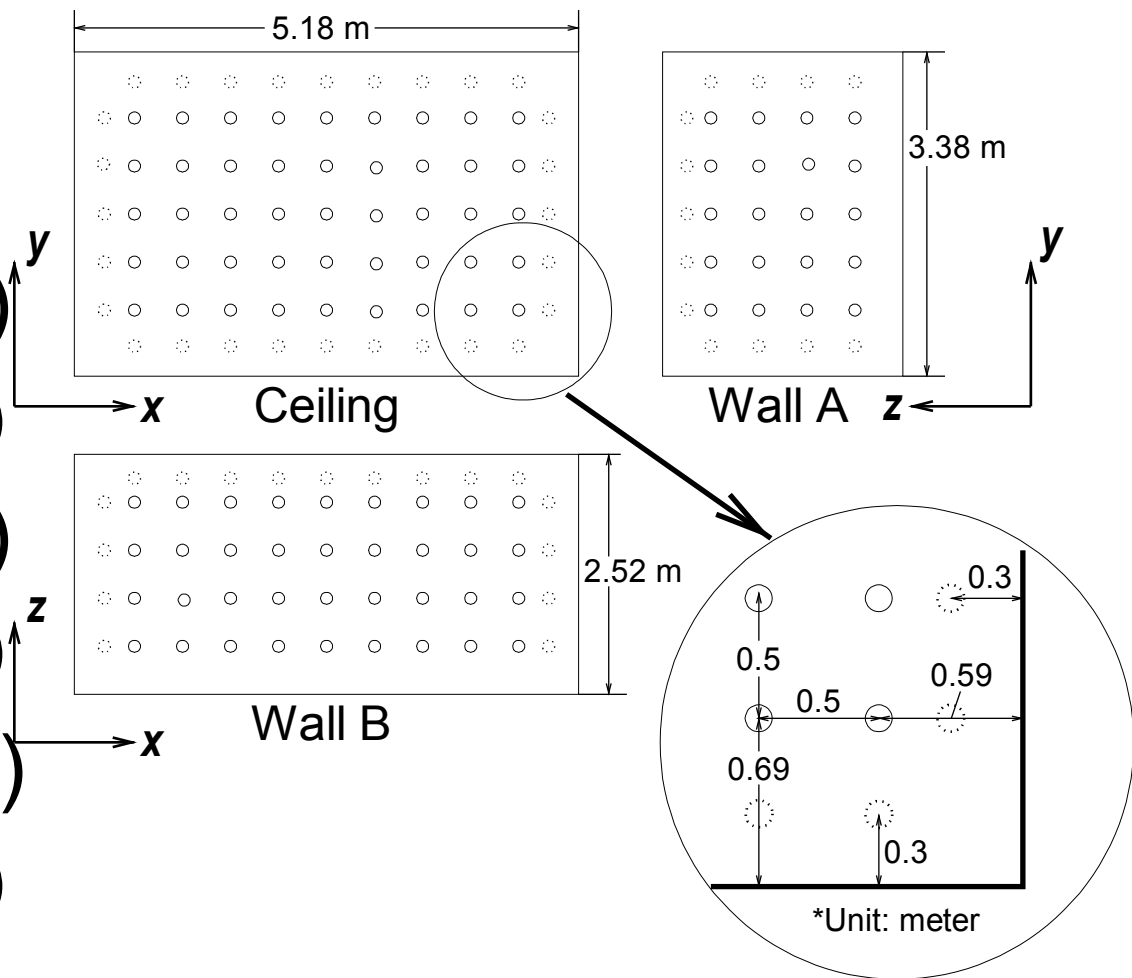
# 包囲型マイクロホンアレイ

- 東北大学電気通信研究所の設備を利用
  - 残響時間
    - 150 ms
  - 防音加工
  - 壁面に吸音マット
  - マイクロホン
    - B&K: Type 4951
    - 157個
  - マイクロホンアンプ
    - B&K: Type 2694
    - 10台(1台あたり16 ch)



# マイクロホンの配置

- マイクロホン間隔...50 cm
- 壁からの距離
  - 30 cm
- 狭壁面(Wall A)
  - 20(=5×4)(2面)
- 広壁面(Wall B)
  - 36(=9×4)(2面)
- 天井面(Ceiling)
  - 45(=9×5)(1面)



# 放射型スピーカアレイ

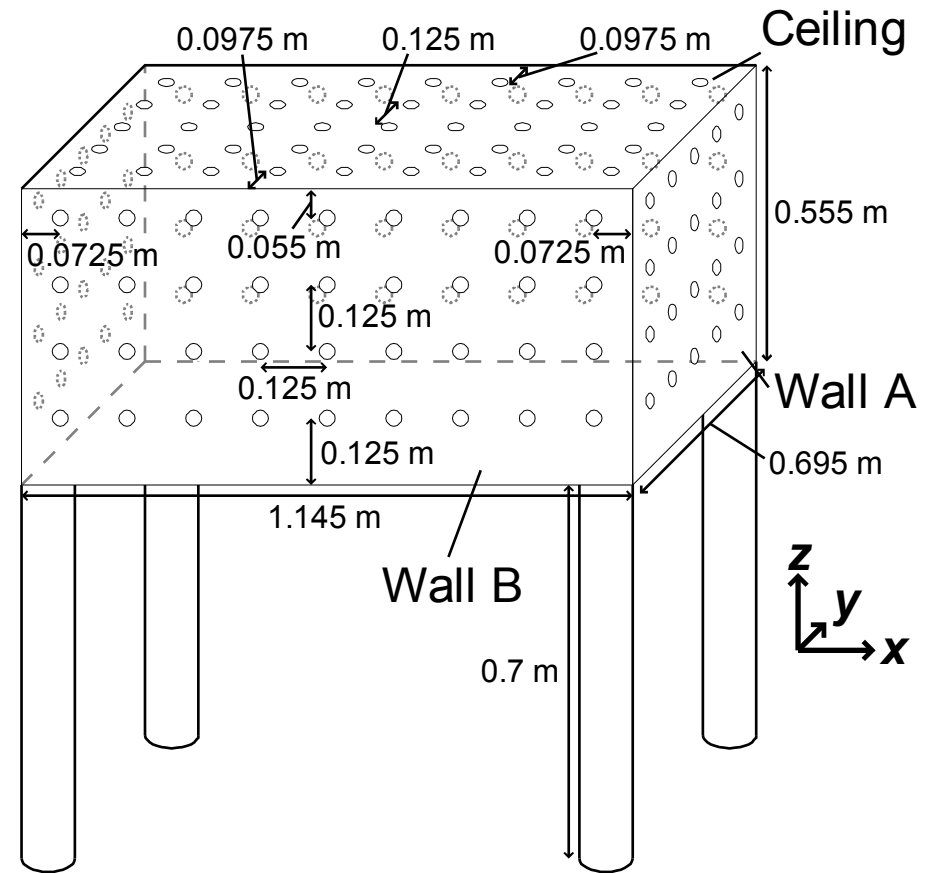
- 新たに製造
  - 大きさ... マイクロホンアレイの1/4
  - スピーカユニット
    - AURASOUND: NSW1-205-8A相当品
    - 157個
    - 指向特性... 外向き
  - スピーカアンプ
    - 特注品
      - 157 ch分





# スピーカユニットの配置

- スピーカユニット間隔...12.5 cm
- 狭壁面(Wall A)
  - 20(=5×4)(2面)
- 広壁面(Wall B)
  - 36(=9×4)(2面)
- 天井面(Ceiling)
  - 45(=9×5)(1面)
- 床上げ
  - 0.7 m



# シミュレーション環境

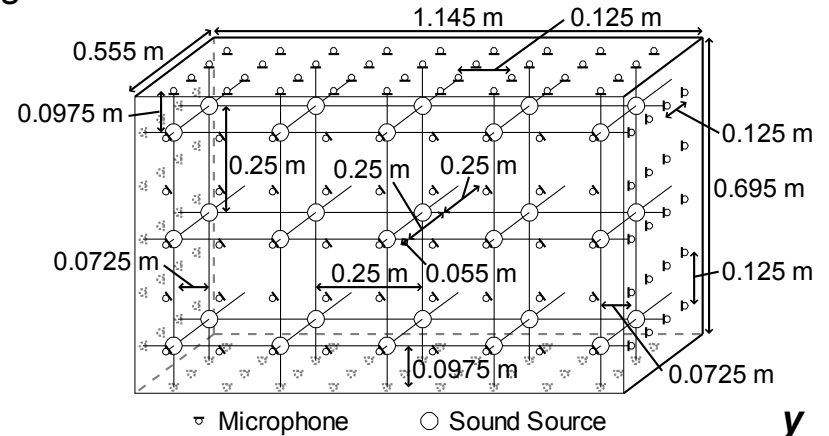
- 原音場

- 自由音場
- マイクホン...157個
- 音源...30個

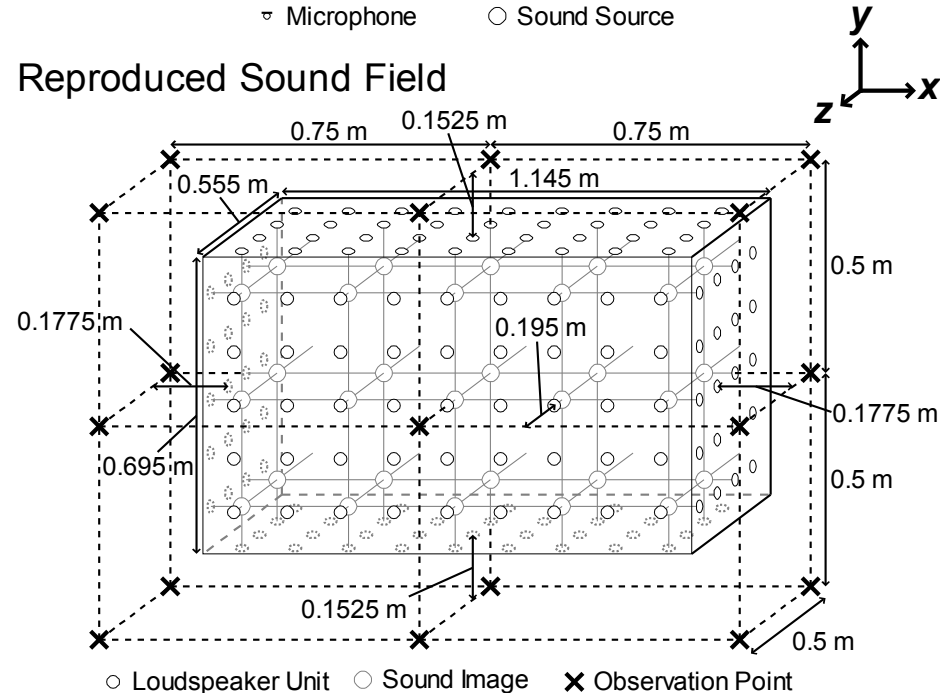
- 再生音場

- 自由音場
- スピーカユニット  
• 157個
- 観測点...17点

Original Sound Field



Reproduced Sound Field



# スピーカで再生する信号の合成

---

- 音源信号  $s(t)$ 
  - オクターブバンドノイズ(中心周波数  $f_{\text{cent}}$ )
- $i$ 番目のスピーカ信号  $x_i(t)$

$$x_i(t) = g_i(t) * s(t) = \frac{1}{|\mathbf{r}_i - \mathbf{r}_0|} s\left(t - \frac{|\mathbf{r}_i - \mathbf{r}_0|}{c}\right)$$

- \*: 畳み込み演算
- $\mathbf{r}_i$ : マイクロホンの位置ベクトル
- $\mathbf{r}_0$ : 音源の位置ベクトル
- $c$ : 音速

# 観測点における音圧の合成

- 観測点 $\mathbf{R}_j$ における瞬時音圧 $p(\mathbf{R}_j, f_{\text{cent}}, t)$

$$\begin{aligned} p(\mathbf{R}_j, f_{\text{cent}}, t) &= \sum_{i=1}^M \frac{D_{si}}{|\mathbf{R}_j - \mathbf{r}_i|} x_i \left( t - \frac{|\mathbf{R}_j - \mathbf{r}_i|}{c} \right) \\ &= \sum_{i=1}^M \frac{D_{si}}{|\mathbf{R}_j - \mathbf{r}_i| |\mathbf{r}_i - \mathbf{r}_0|} s \left( t - \frac{|\mathbf{R}_j - \mathbf{r}_i| + |\mathbf{r}_i - \mathbf{r}_0|}{c} \right) \end{aligned}$$

- $M$ : スピーカユニットの総数
- $D_{si}$ : スピーカユニットの放射指向特性

# 音響インテンシティの算出

- 音響インテンシティベクトルの方向

- 単一音源の到来方向に相当

$$\mathbf{I}(\mathbf{R}_j, f_{\text{cent}}) = \left\{ I_x(\mathbf{R}_j, f_{\text{cent}}), I_y(\mathbf{R}_j, f_{\text{cent}}), I_z(\mathbf{R}_j, f_{\text{cent}}) \right\}^T$$

- クロススペクトル法により算出

- 6点の音圧から3方向を算出

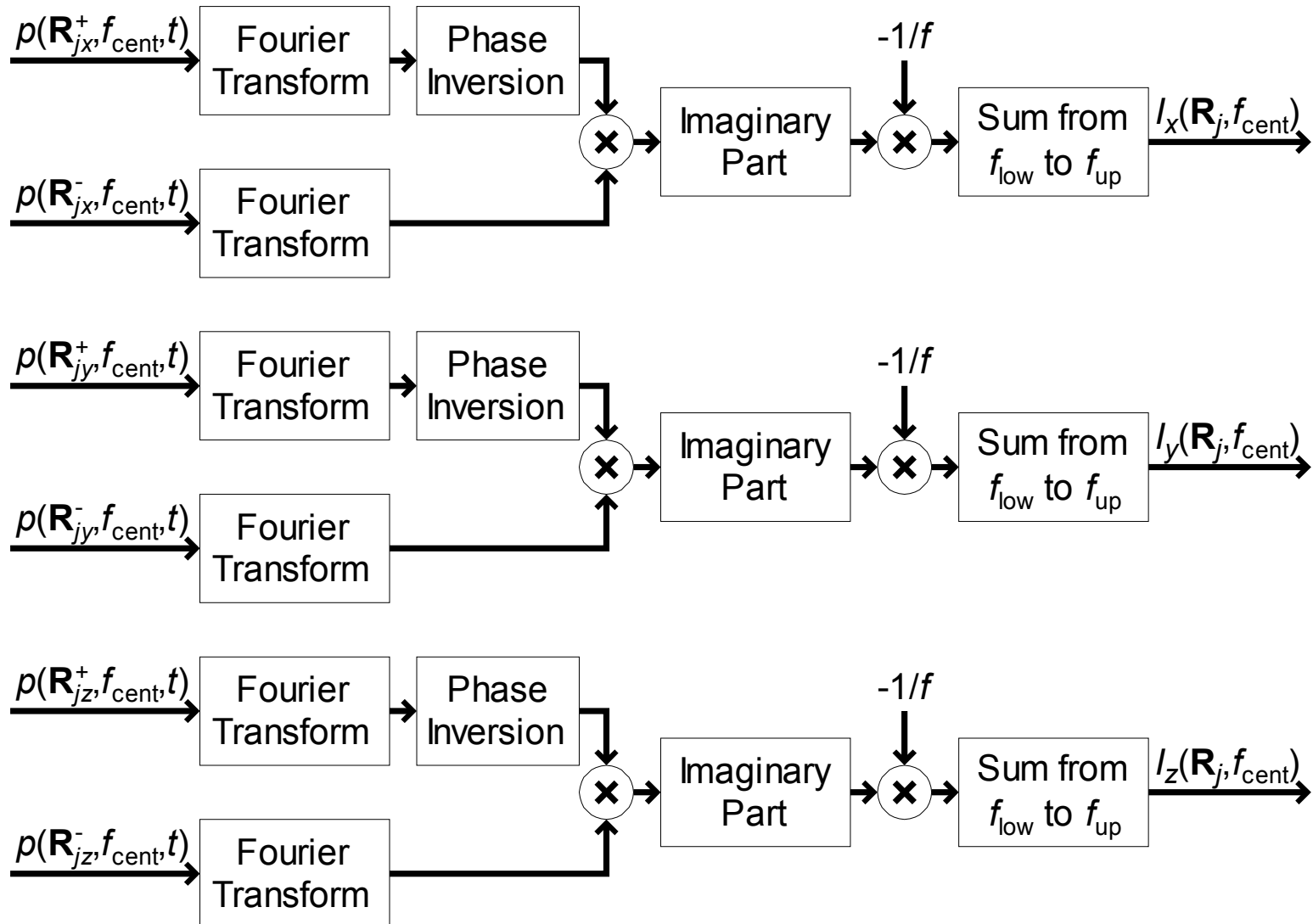
- 音圧:  $p(\mathbf{R}_{jx}^+, f_{\text{cent}}, t)$ ,  $p(\mathbf{R}_{jx}^-, f_{\text{cent}}, t)$ ,  $p(\mathbf{R}_{jy}^+, f_{\text{cent}}, t)$ ,  
 $p(\mathbf{R}_{jy}^-, f_{\text{cent}}, t)$ ,  $p(\mathbf{R}_{jz}^+, f_{\text{cent}}, t)$ ,  $p(\mathbf{R}_{jz}^-, f_{\text{cent}}, t)$ ,

$$\mathbf{R}_{jx}^{\pm} = \mathbf{R}_j \pm (\Delta, 0, 0)^T$$

$$\mathbf{R}_{jy}^{\pm} = \mathbf{R}_j \pm (0, \Delta, 0)^T \quad \Delta = 0.001 \text{ m}$$

$$\mathbf{R}_{jz}^{\pm} = \mathbf{R}_j \pm (0, 0, \Delta)^T$$

# 音響インテンシティ算出のブロック図

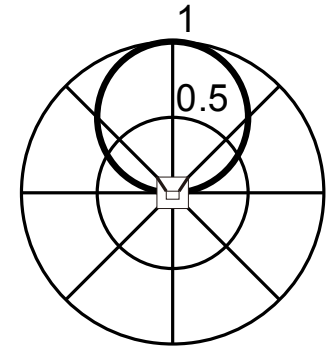
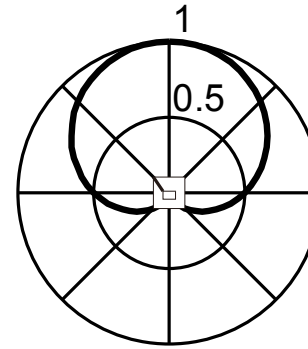
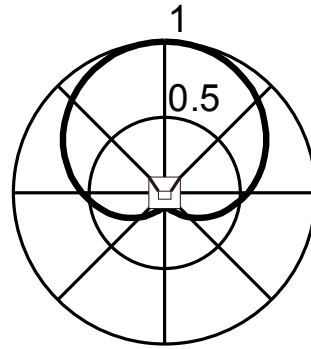
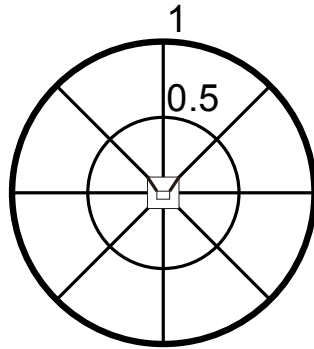


# パラメータ条件

中心周波数( $f_{\text{cent}}$ )	250, 500, 1000 Hz
下限周波数( $f_{\text{low}}$ )	$f_{\text{cent}} \div \text{sqrt}(2)$
上限周波数( $f_{\text{up}}$ )	$f_{\text{cent}} \times \text{sqrt}(2)$
音速( $c$ )	340 m/s
スピーカユニットの数( $M$ )	157
スピーカユニットの放射指向特性( $D_{si}$ )	Omnidirectional, Decay20dB, Unidirectional, Shotgun

# スピーカユニットの放射指向特性

Omnidirectional Decay 20dB    Unidirectional    Shotgun



(Omnidirectional)

$$D_{si} = 1$$

(Decay 20dB)

$$D_{si} = 0.55 + 0.45 \cos \theta_{si}$$

(Unidirectional)

$$D_{si} = \frac{1 + \cos \theta_{si}}{2}$$

$$\cos \theta_{si} = \frac{\mathbf{n}_{si} \cdot (\mathbf{R}_j - \mathbf{r}_i)}{|\mathbf{n}_{si}| |\mathbf{R}_j - \mathbf{r}_i|}$$

(Shotgun)

$$D_{si} = \begin{cases} \cos \theta_{si} & (\theta_{si} \leq 90^\circ) \\ 0 & (\theta_{si} > 90^\circ) \end{cases}$$



# 音像位置の推定

- 推定した音像位置  $\mathbf{r}_E$

$$\mathbf{r}_E = \frac{1}{FN} \sum_{f_{\text{cent}} \in \{250, 500, 1000\}} \sum_{j=1}^N \left\{ \mathbf{R}_j - \frac{\mathbf{I}(\mathbf{R}_j, f_{\text{cent}})}{\rho(\mathbf{R}_j, f_{\text{cent}})} \right\}$$

–  $\mathbf{R}_j$ :  $j$ 番目の観測点の位置ベクトル

–  $\mathbf{I}(\mathbf{R}_j, f_{\text{cent}})$ :  $\mathbf{R}_j$ における音響インテンシティ

–  $\rho(\mathbf{R}_j, f_{\text{cent}})$ :  $\mathbf{R}_j$ における平均二乗音圧

–  $T$ : 周期  $\rho(\mathbf{R}_j, f_{\text{cent}}) = \sqrt{\frac{1}{T} \int_0^T \{p(\mathbf{R}_j, f_{\text{cent}}, t)\}^2 dt}$

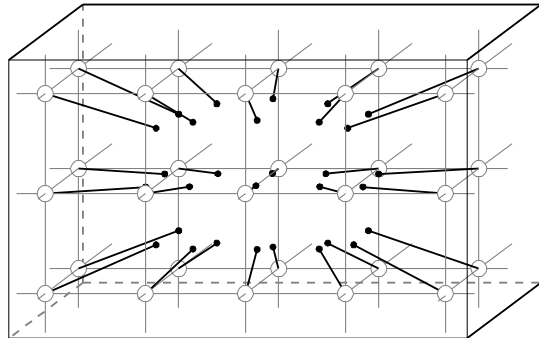
–  $F(=3)$ : オクターブバンドノイズの数

–  $N$ : 観測点の数(全て, 音源に近接した4点)

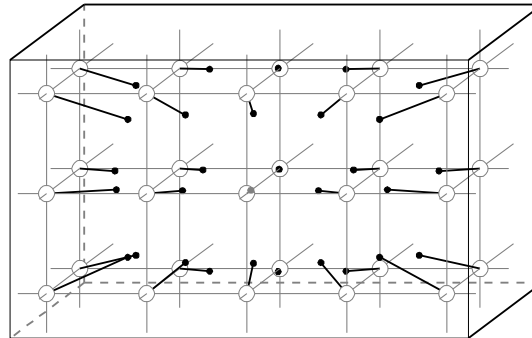
# 推定した音像位置(全ての観測点)

- 放射指向性が鋭いほど正しく推定
  - 放射指向性が鋭いほど，聴取者はどの位置においても常に正しく音像を定位

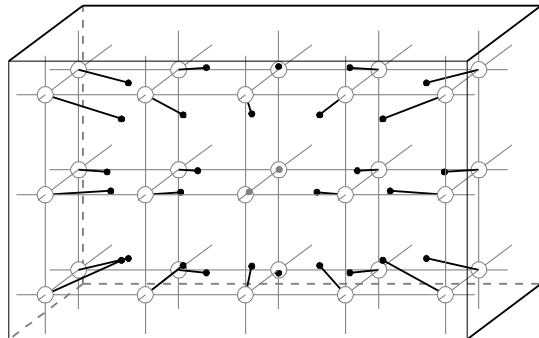
Omnidirectional



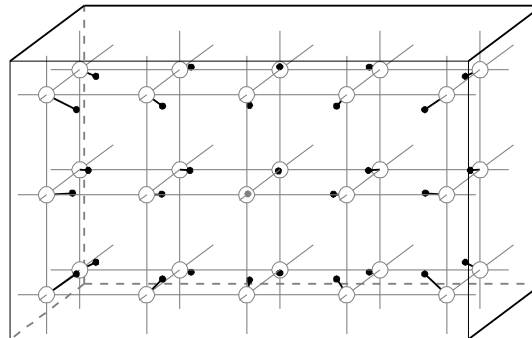
Decay 20dB



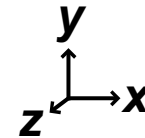
Unidirectional



Shotgun



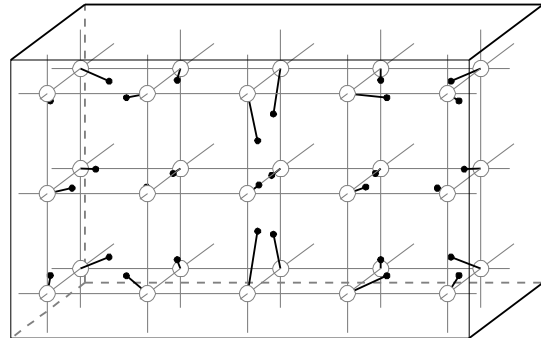
- Input Sound Image
- Estimated Sound Image



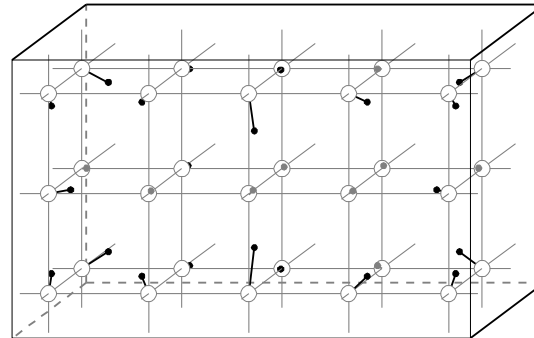
# 推定した音像位置(近接した四観測点)

- どの放射指向性でも大体正しく推定
  - 聴取者は聴取位置で止まって聞けば, 聴取位置の近くにある音像を常に正しく定位

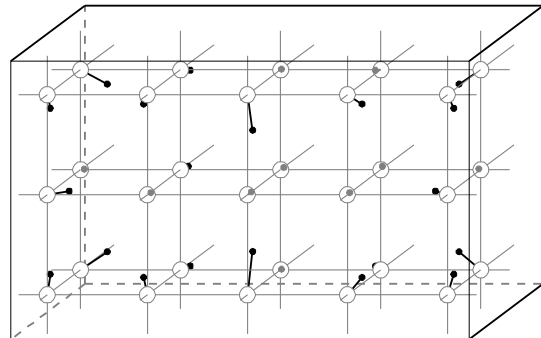
Omnidirectional



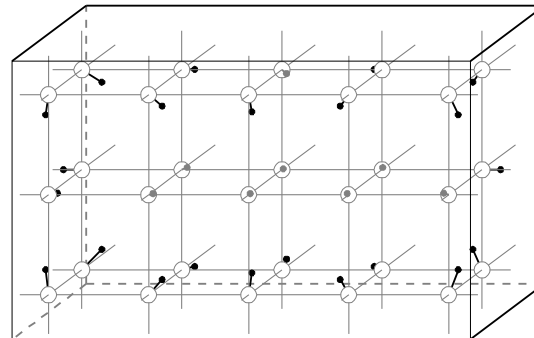
Decay 20dB



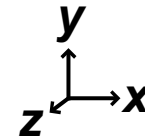
Unidirectional



Shotgun

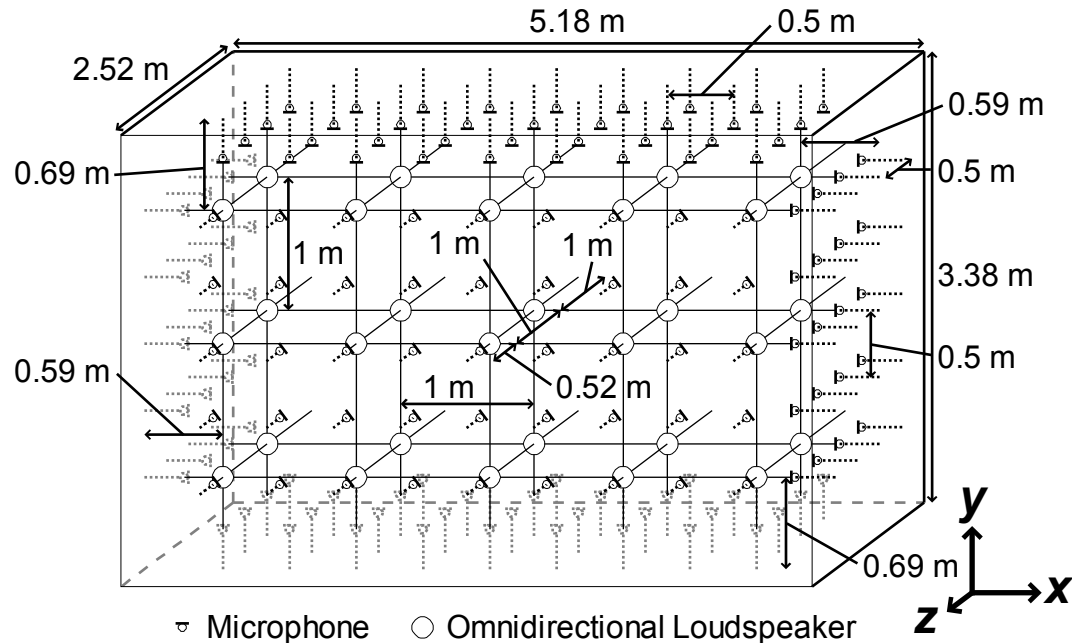


- Input Sound Image
- Estimated Sound Image



# 室内インパルス応答の測定

- 30ヶ所に無指向性スピーカを配置
- 残響時間
  - 150 ms
- 室温
  - 20°C
- 暗騒音レベル
  - 18.4 dB(A)
- 音圧レベル
  - 無指向性スピーカより1 mで85.6 dB(A)



# 測定条件

---

- TSP信号
  - 標本化周波数...48 kHz
  - 量子化ビット...16 bits
  - サンプル長...65536 samples
- FIRフィルタ
  - タップ長...14400点
  - 同期加算...16回
  - 標本化周波数...192 kHzとみなす
    - マイクロホンとスピーカユニットの位置を同じにするため

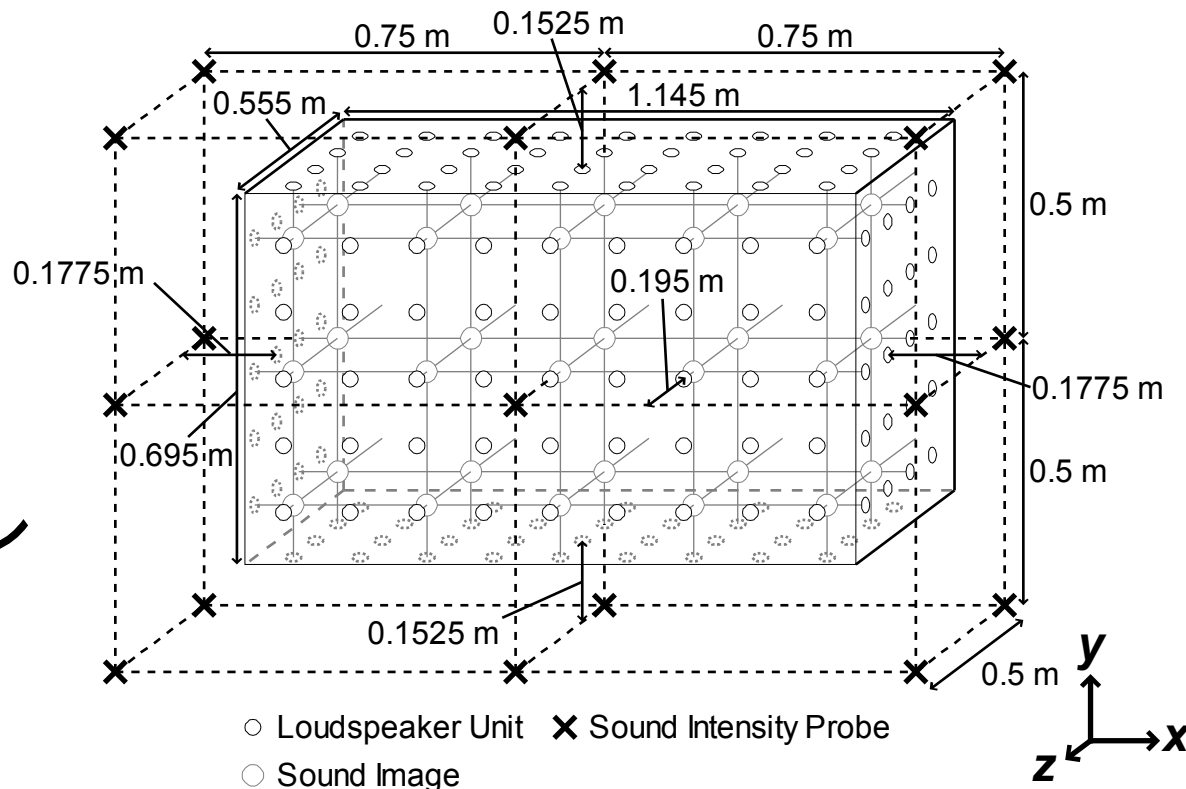
# 157チャンネル信号の作成

---

- 音源信号
  - オクターブバンドノイズ
    - 中心周波数...250, 500, 1000 Hz
  - 標本化周波数...48 kHz
  - 量子化ビット...16 bits
  - 長さ...10 s
  - フェードイン&アウト時間...1 ms
- 室内インパルス応答
  - 畳み込み前に48 kHzにリサンプリング処理

# 音響測定

- スピーカアレイ外部の17ヶ所で音響測定
- 残響時間
  - 180 ms
- 室温
  - 22°C
- 暗騒音レベル
  - 22 dB(A)
- 音圧レベル
  - スピーカアレイ中心音像から1 mで71 dB(A)



# 音像位置の推定

- 推定した音像位置  $\mathbf{r}_E$

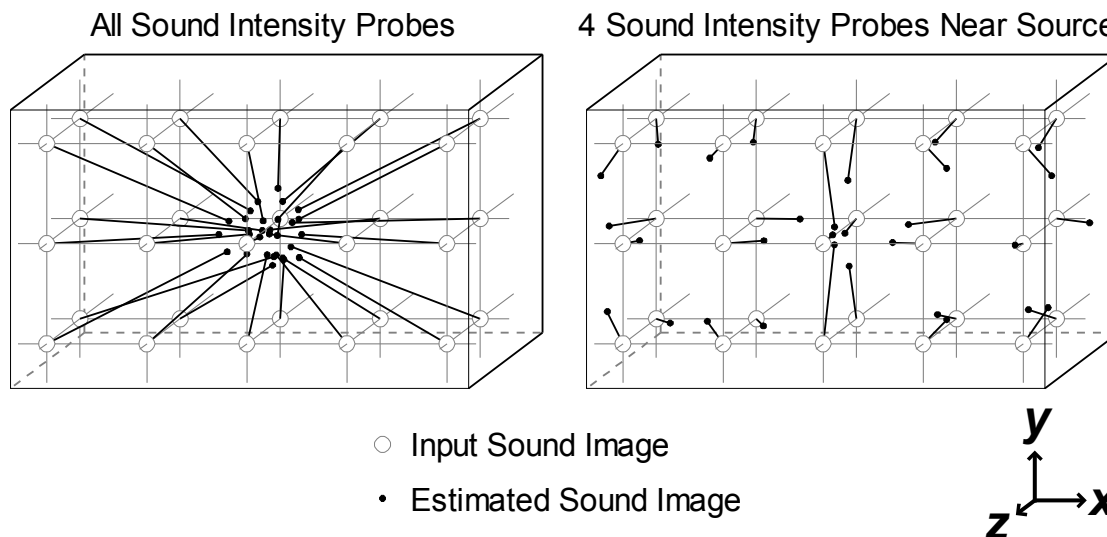
$$\mathbf{r}_E = \frac{1}{FN} \sum_{f_{\text{cent}} \in \{250, 500, 1000\}} \sum_{j=1}^N \left\{ \mathbf{R}_j - \frac{\mathbf{I}(\mathbf{R}_j, f_{\text{cent}})}{\rho(\mathbf{R}_j, f_{\text{cent}})} \right\}$$

- $\mathbf{R}_j$ :  $j$ 番目の観測点の位置ベクトル
- $\mathbf{I}(\mathbf{R}_j, f_{\text{cent}})$ :  $\mathbf{R}_j$ における音響インテンシティ
- $\rho(\mathbf{R}_j, f_{\text{cent}})$ :  $\mathbf{R}_j$ における音圧レベル
- $F(=3)$ : オクターブバンドノイズの数
- $N$ : 観測点の数(全て, 音源に近接した4点)



# 推定した音像位置

- 近接した4点
  - スピーカユニットの放射指向性は無指向性
- 全ての点
  - 測定誤差が加味
  - 放射指向性がより鋭ければより良い性能に



# まとめ

---

- 指向性スピーカと波面合成法を用いた近接三次元音場再生システムを評価
  - 計算機シミュレーションによる音像定位性能
    - 放射指向性が鋭いほどどの場所でも定位できる
    - 鋭くなくても聴取位置近くの音像は定位できる
  - 音響測定による音像定位性能
    - 放射指向性は全体的に無指向性
    - 聴取者は聴取位置で止まって聞けば、聴取位置の近くにある音像を常に正しく定位
- 今後の課題
  - 放射指向性をより鋭くし、音像定位性能向上