

マイクロホンアレーを用いた 自由聴点音場再生システムに 関する理論的検討

名古屋大学 大学院 情報科学研究科

水野 渉

木村敏幸 西野隆典 武田一哉

はじめに

◆収録環境の再現

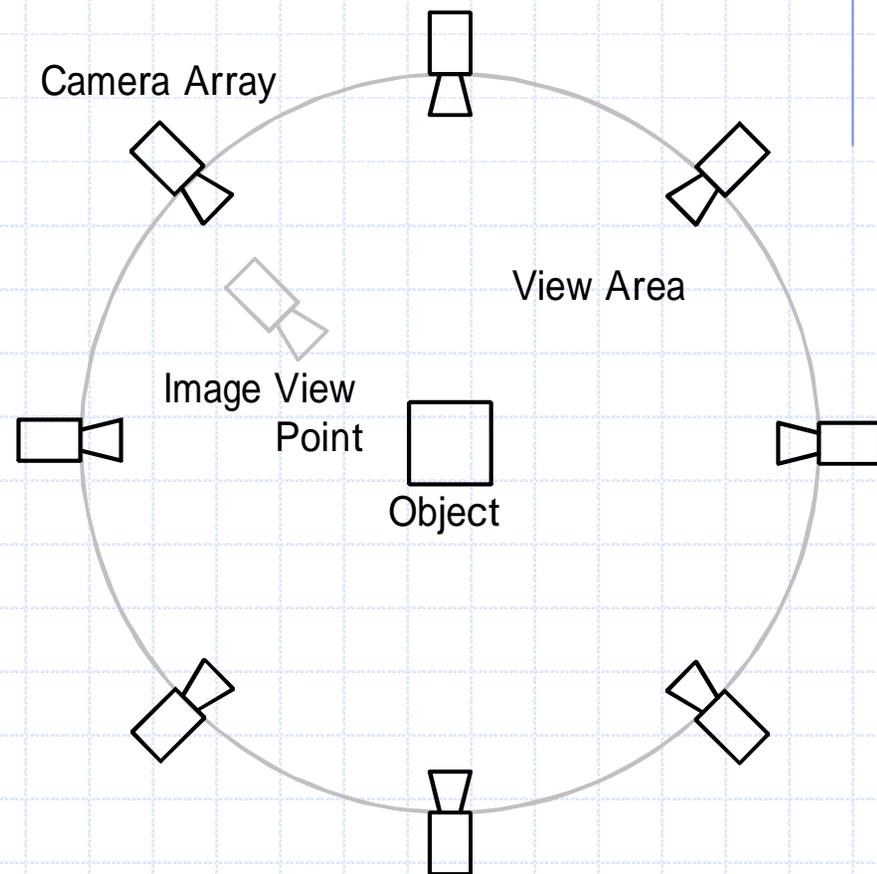
- 異なる場所に居ながらコンサートホールなどの臨場感
- 立体音響提示
- 任意視点の映像の提示

◆必要となる技術

- 音場再現システム
- 映像生成システム

音場再現と映像生成の統合

- 自由視点テレビシステム
 - ◆ 対象物をカメラアレーで囲む
 - ◆ 収録した映像から任意視点の映像を生成
- 任意視点の聴覚情報が必要
 - ◆ 音源の位置を限定しない
- 聴取領域の制御が必要
- 波面合成法の利用

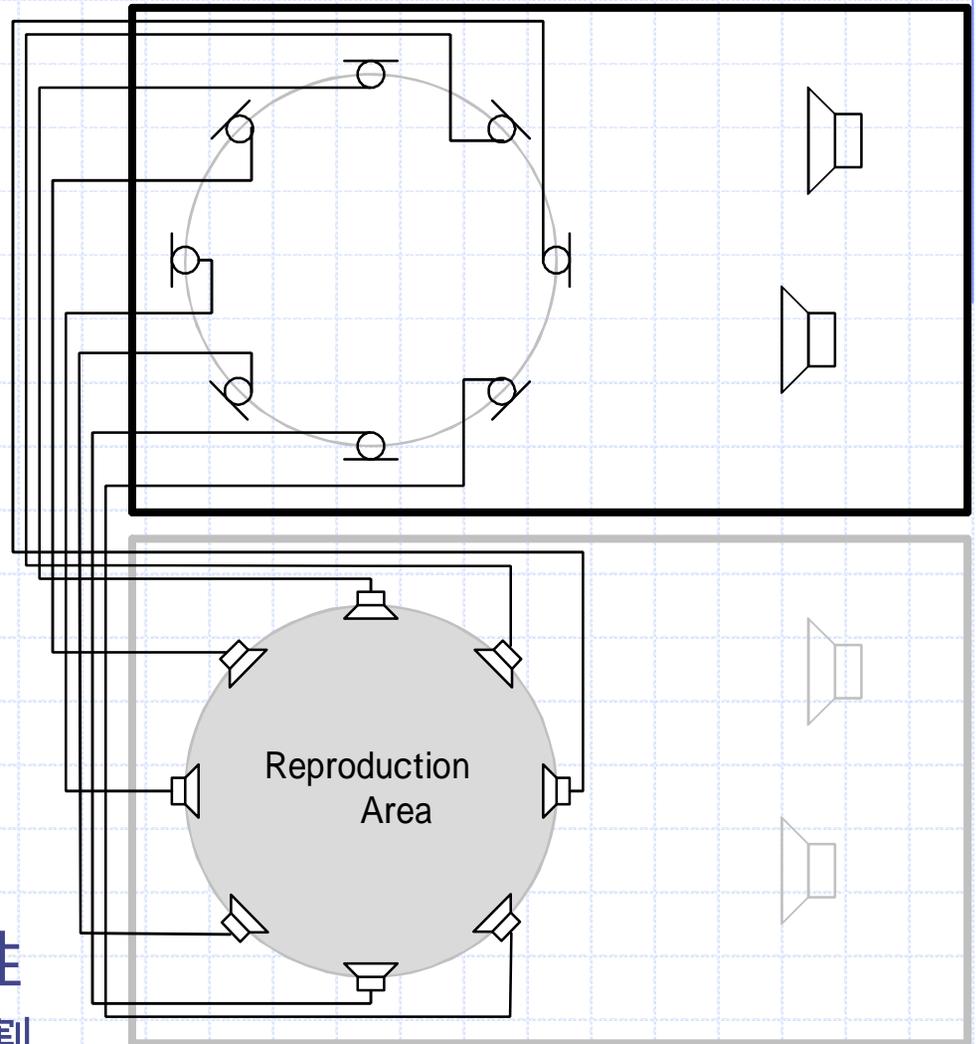


波面合成法

- マイクロホンアレーを用いて原音場を収録
- 同一に配置したスピーカアレーを用いて別の場所で収録音を再生
- Huygensの原理によって音場を再現

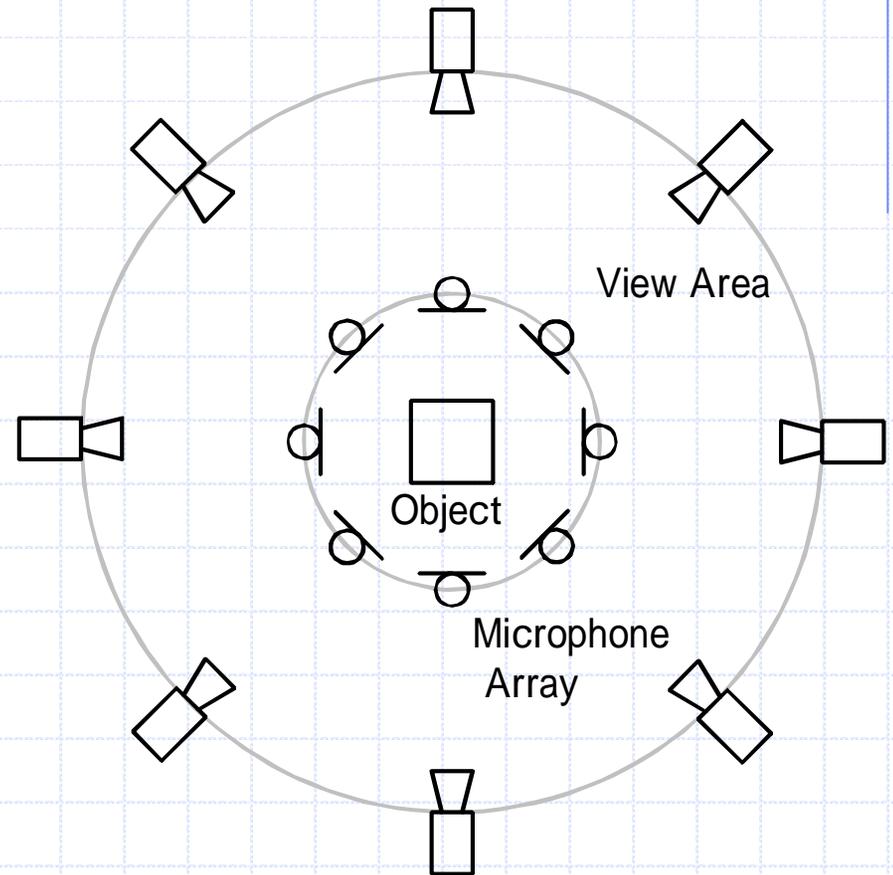


- マイクとスピーカの指向性
 - ◆ 音源領域と再現領域を分割



システム統合

- 音源の周囲をマイクロホンアレーで囲む

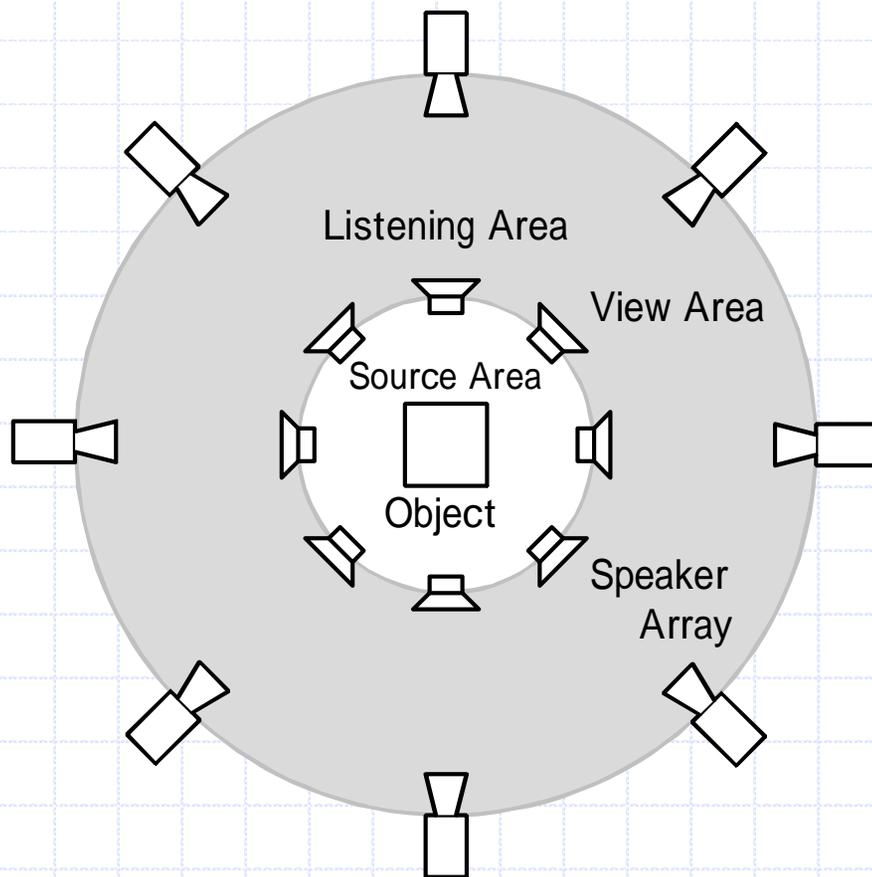


システム統合

- 音源の周囲をマイクロホンアレーで囲む



- カメラアレーとマイクロホンアレーの間の領域が再現領域



システム統合の問題点

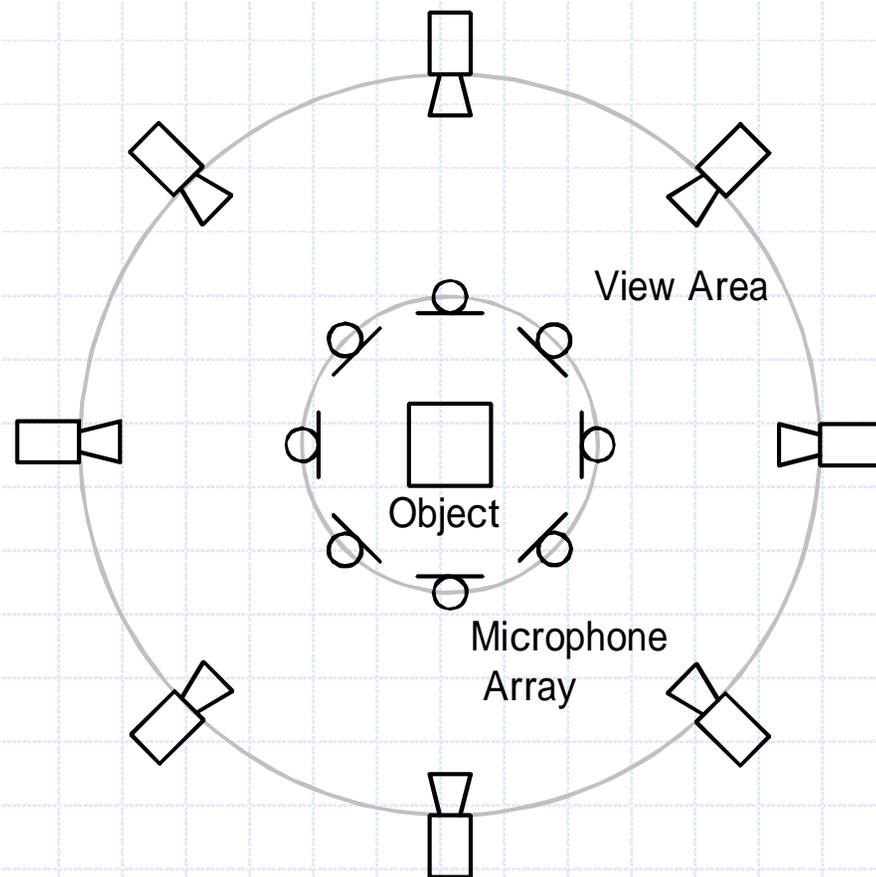
- 音源の周囲をマイクロホンアレーで囲む



- カメラアレーとマイクロホンアレーの間の領域が再現領域



- 収録映像にマイクロホンが映り込む事が問題

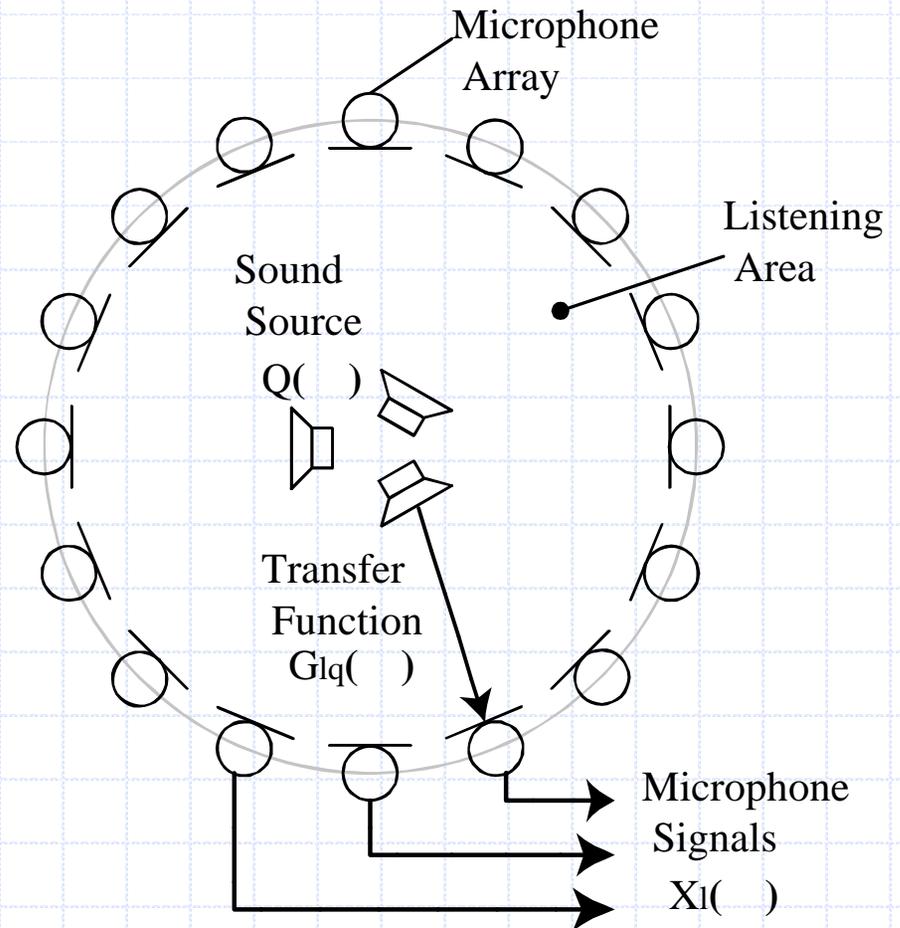


本研究の目的

- ◆ 自由視点テレビシステムにおける聴覚情報の再現
 - 波面合成の原理による聴取領域内の音場再現
 - ◆ ヘッドホンを用いた提示システム
 - カメラアレーの領域内にマイクロホンアレーは設置できない
- 
- 領域内部のマイクロホンアレー信号の推定
 - ◆ カメラアレーと同一位置に設置したマイクロホンアレーを使用
 - 仮想音源による再現
 - ◆ 推定したマイクロホン信号を出力
 - 検討課題
 - ◆ 必要な仮想音源数、マイクロホン数、再現精度

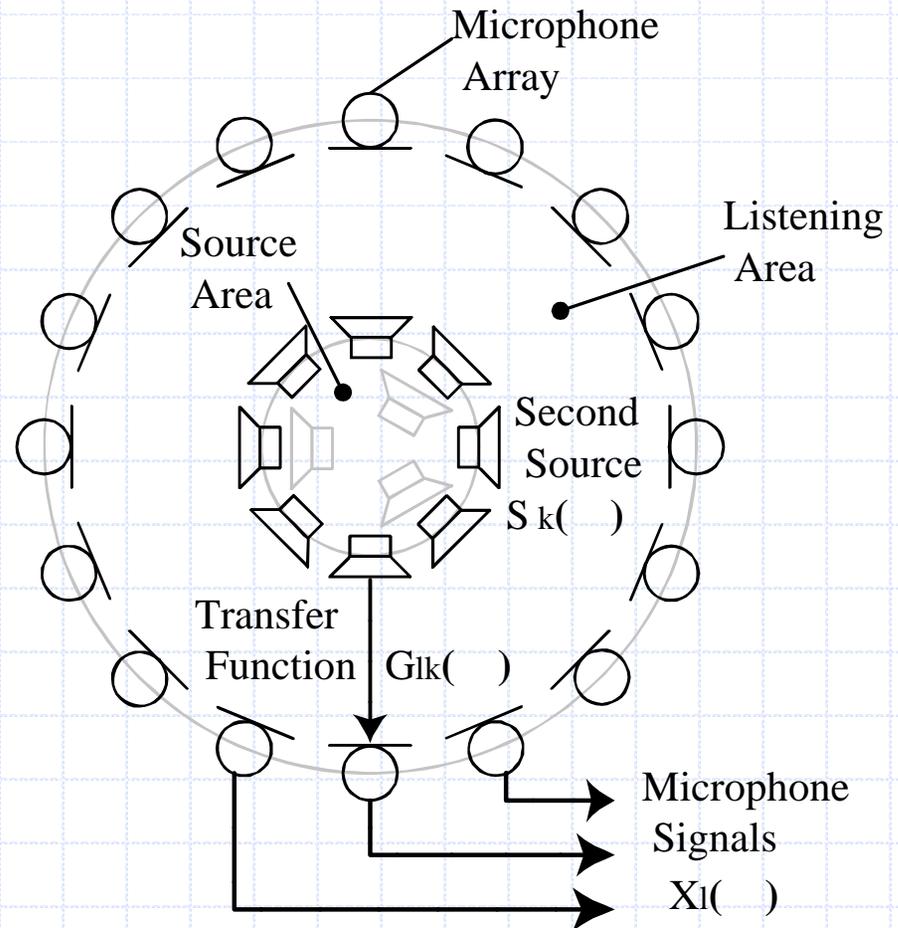
提案手法 - 実音源信号による収録

- マイクロホン信号 X は実音源信号によって得られる



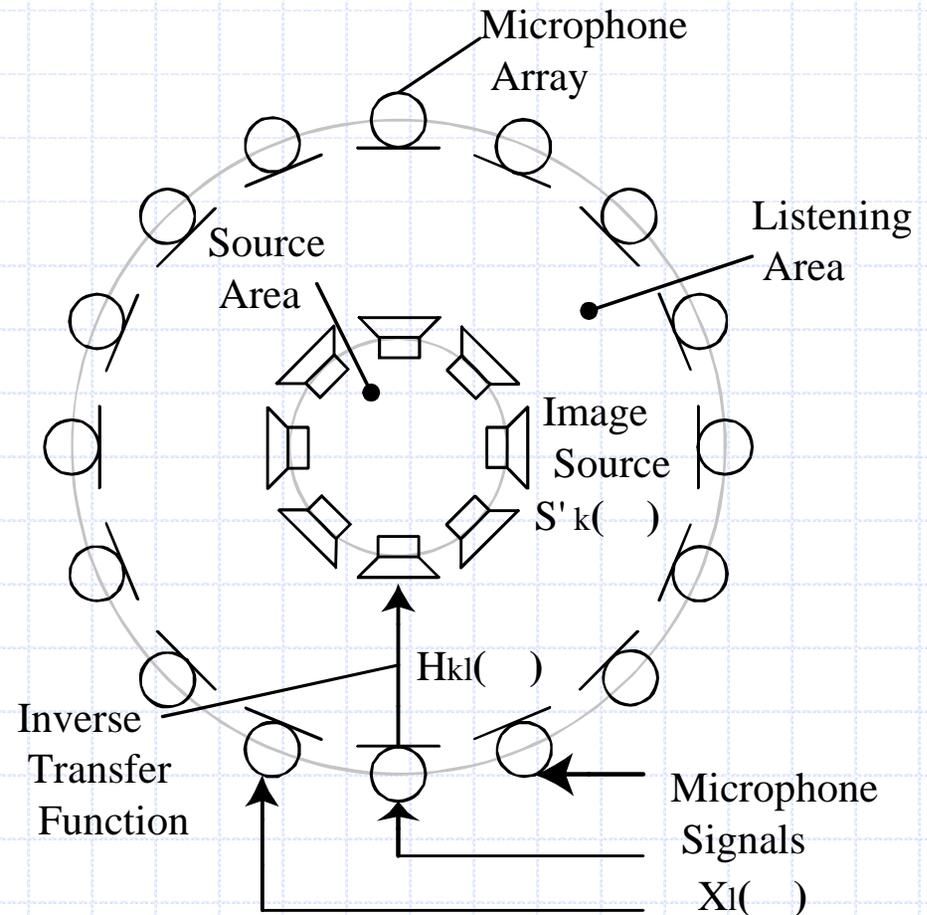
提案手法 - 2次音源信号の仮定

- マイクロホン信号 X は実音源信号によって得られる
- ↓
- マイクロホン信号 X は2次音源信号 S による波面合成によって得られたと仮定

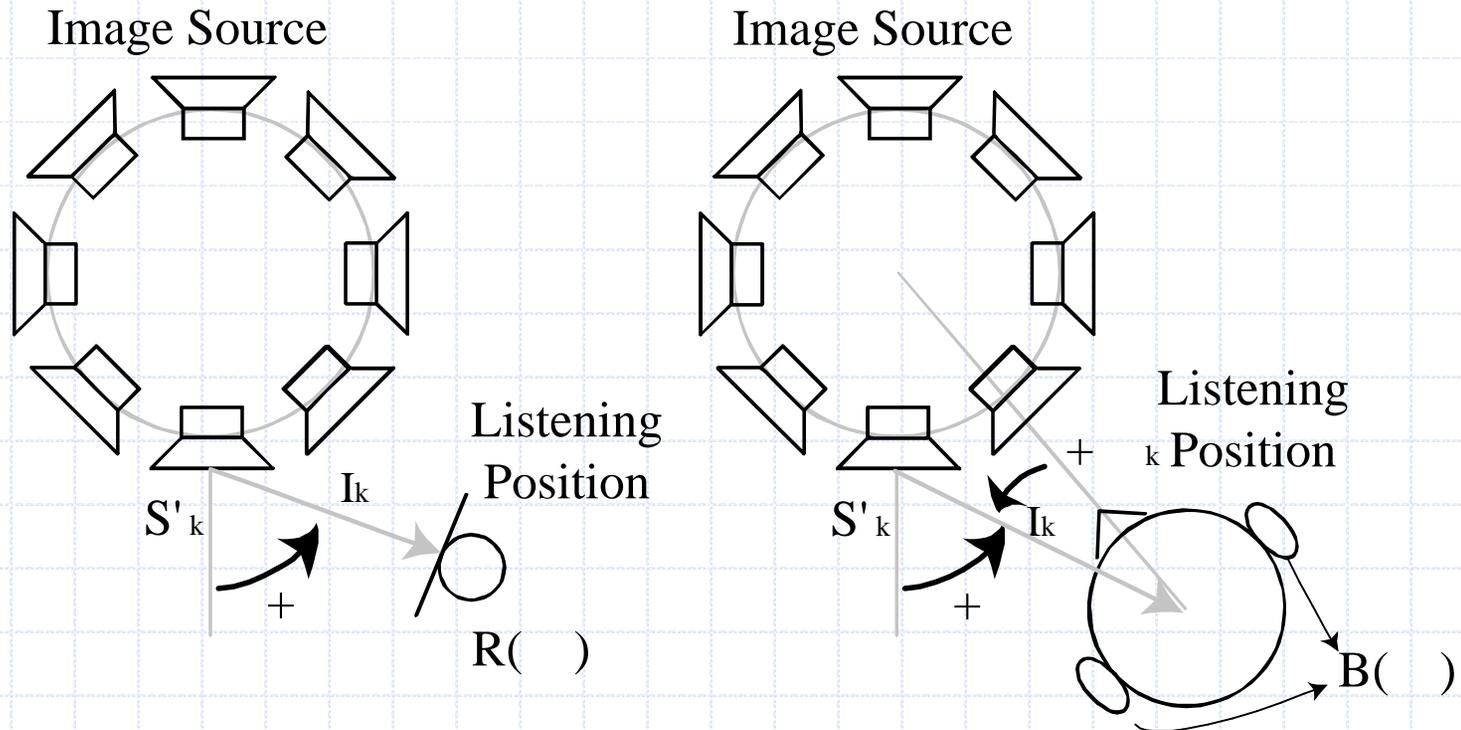


提案手法 - 仮想音源信号の算出

- マイクロホン信号 X は実音源信号によって得られる
- マイクロホン信号 X は2次音源信号 S による波面合成によって得られたと仮定
- 逆伝達関数 H を用いて仮想音源信号 S' の算出



提案手法 - 仮想音源による音場再現



◆ 仮想音源信号の決定

- 任意の聴取位置の信号が再現可能
- 頭部伝達関数を使い、ヘッドホン提示するシステム

音場再現実験

◆ 提案手法の実装

- マイクロホン信号の生成
- 逆伝達特性の算出
- 仮想音源信号の推定

◆ 客観評価

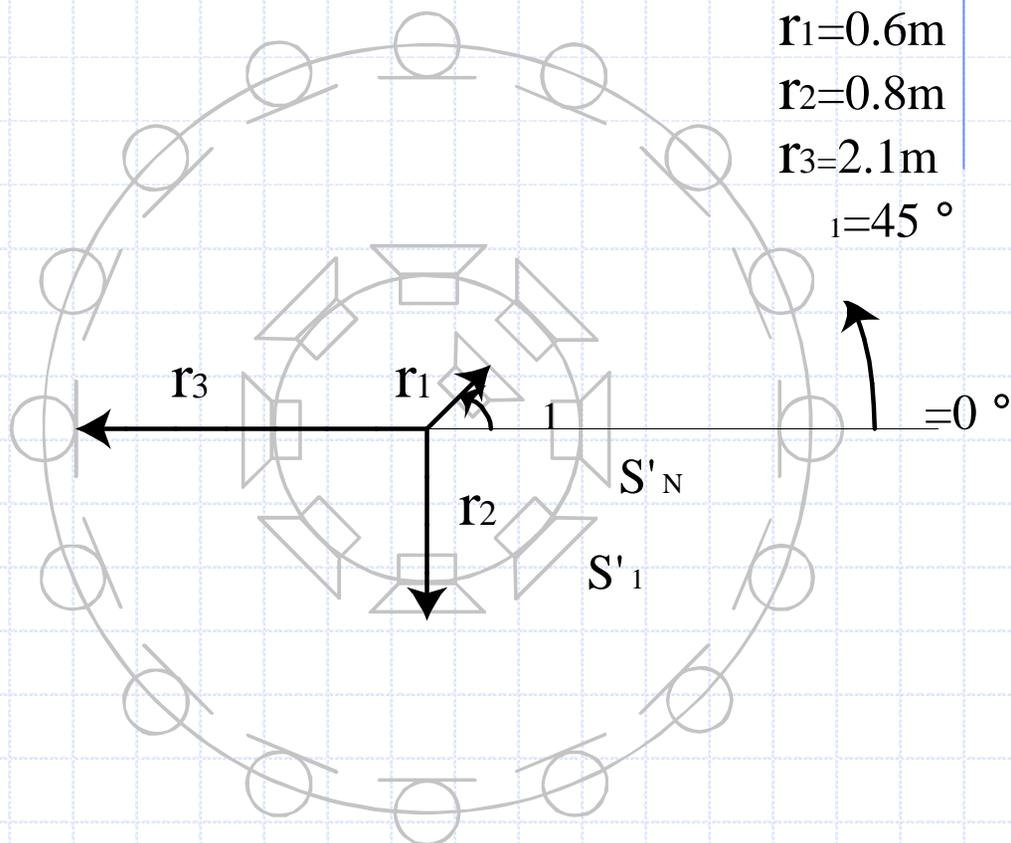
- 合成した音圧信号と実音源が作る信号の比較

◆ 主観評価

- バイノーラル信号の合成
- 実音源が作るバイノーラル信号と比較
 - ◆ ヘッドホン受聴を想定されたシステム

実験時の機器の配置

- 同心円上に配置
- 実音源
 - ◆ 方位角 45°
 - ◆ 中心からの距離 0.6m
- 仮想音源
 - ◆ 方位角 0° 位置に S'_N
 - ◆ 中心からの距離 0.8m 上に均等に配置
 - ◆ 時計回りに配置
- マイクロホン
 - ◆ 中心からの距離 2.1m 上に均等に配置



実験条件

- 伝達関数G
 - ◆ 距離から算出される遅延と減衰を持つデルタ関数
- FIRバンドパスフィルタ
 - ◆ 遅延付加と帯域制限250Hz-13.3kHz
- サンプリング周波数 32kHz
- 音速 340m/s
- 仮想音源の指向性
 - ◆ 外向きの指向性

$$c_k(\Delta) = \begin{cases} \cos(\Delta) & -90^\circ < \Delta < 90^\circ \\ 0 & -180^\circ < \Delta \leq -90^\circ, 90^\circ \leq \Delta \leq 180^\circ \end{cases}$$

推定した信号の評価

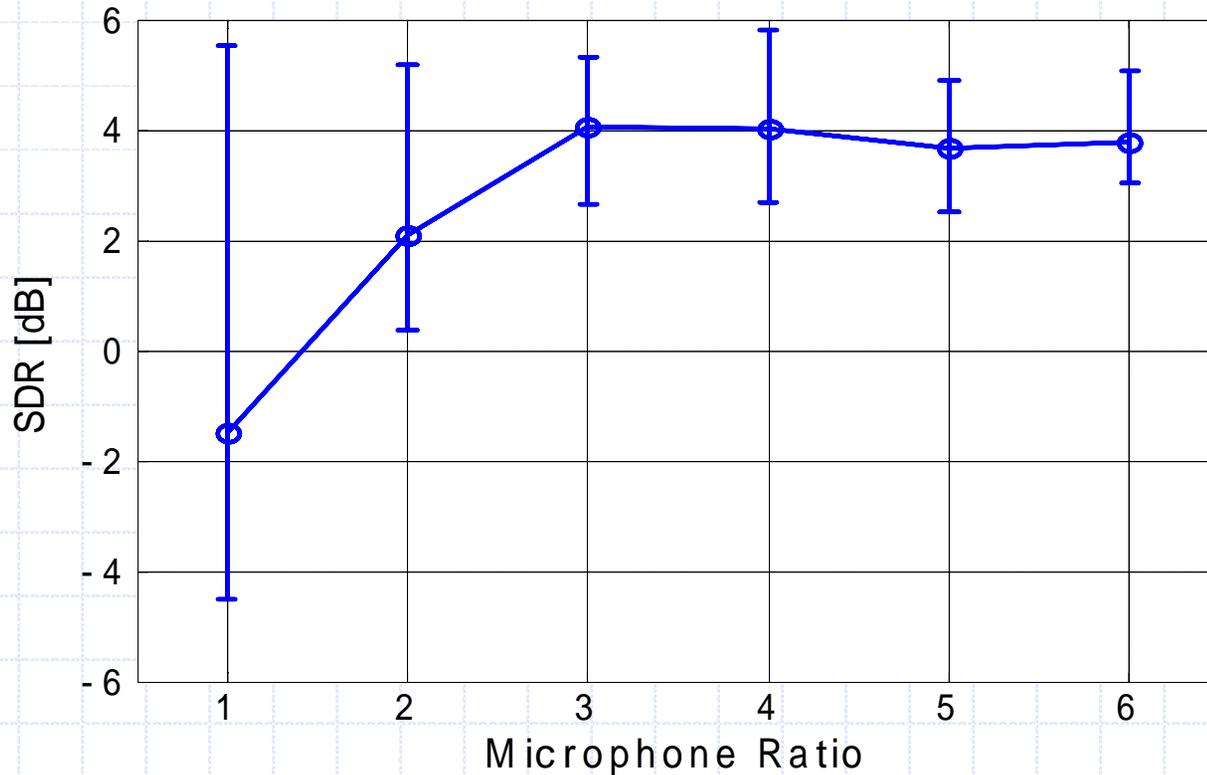
◆ 波形の推定精度の調査

- SDR: Signal to Deviation Ratio

$$\text{SDR} = 10 \log_{10} \frac{\sum_n s^2(n)}{\sum_n [s(n) - s'(n)]^2} \quad [\text{dB}]$$

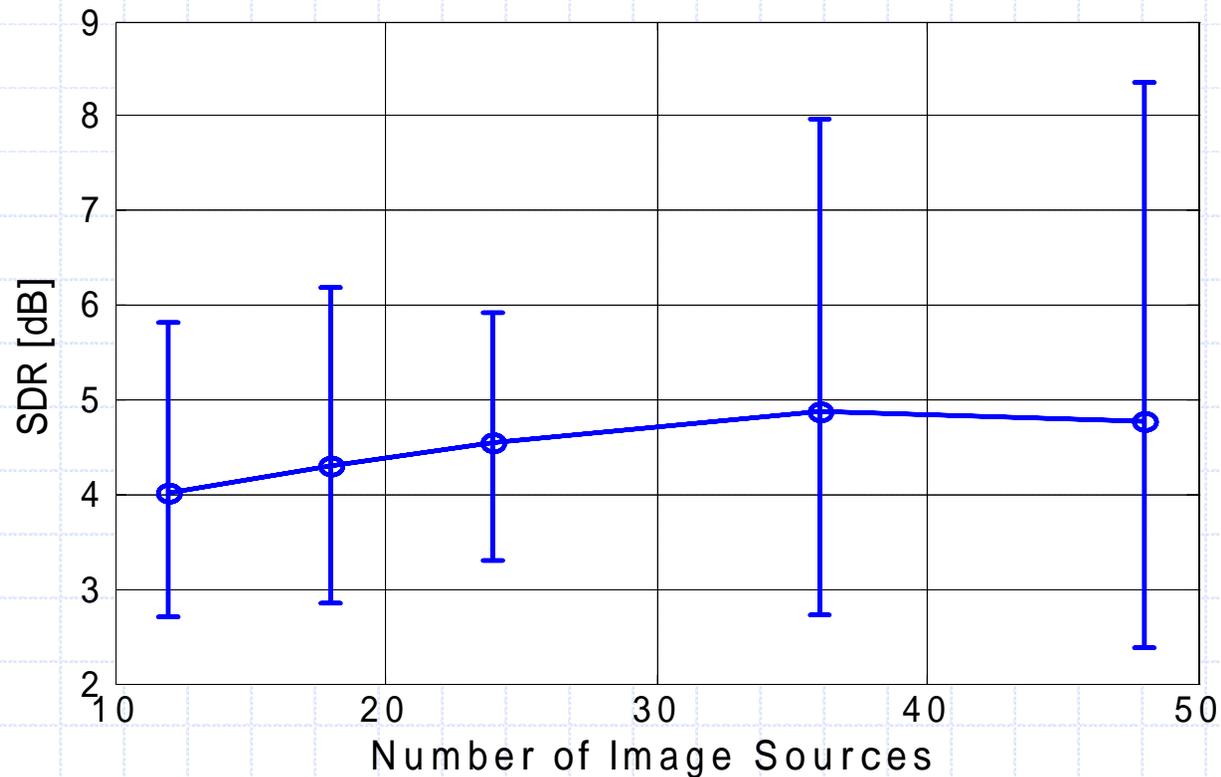
- $s(n)$
 - ◆ 実音源が作る2次音源信号
- $s'(n)$
 - ◆ マイクロホンから推定した仮想音源信号

仮想音源信号の推定とマイク数の比



- 仮想音源を12に固定
- マイクロホンは仮想音源の3倍から4倍必要

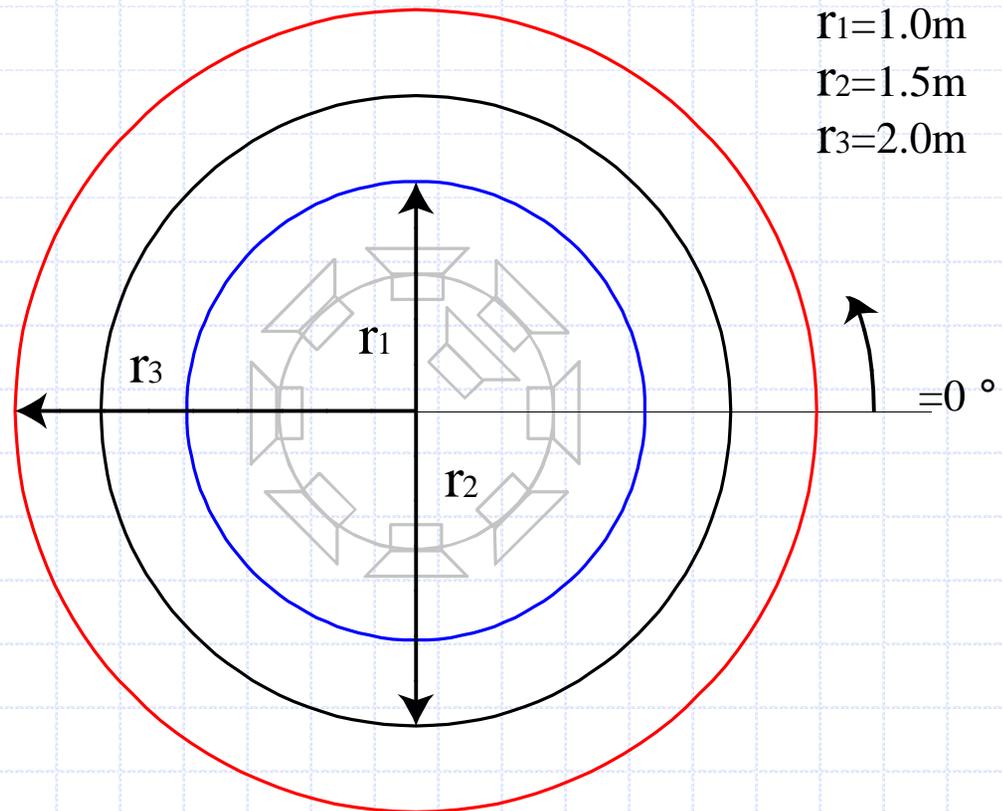
仮想音源信号の推定精度



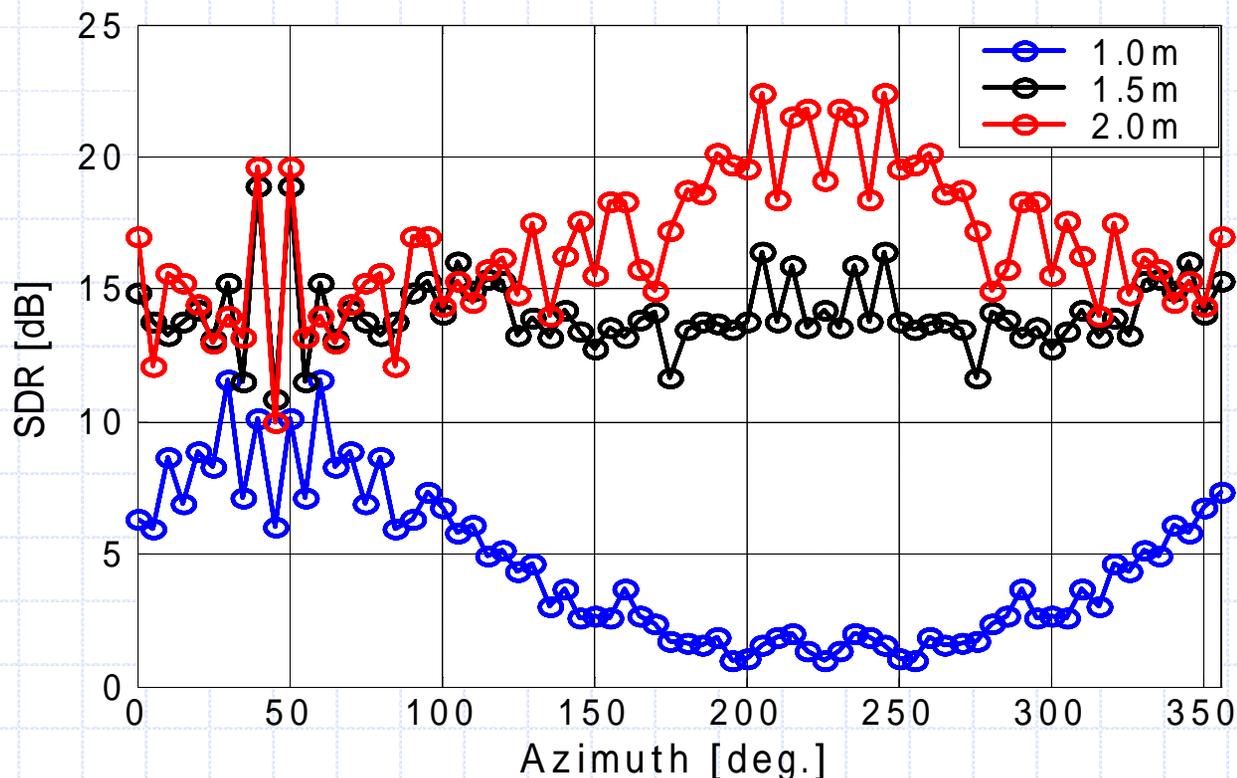
- マイクロホンは仮想音源の4倍に固定
- 仮想音源の平均SDRは4dBから5dB

方位角と距離による再現精度

- 仮想音源数
 - ◆ $N=36$ に固定
- 半径を変化させる
 - ◆ 1.0m
 - ◆ 1.5m
 - ◆ 2.0m
- 円周上で再現
 - ◆ 方位各 5° 刻み



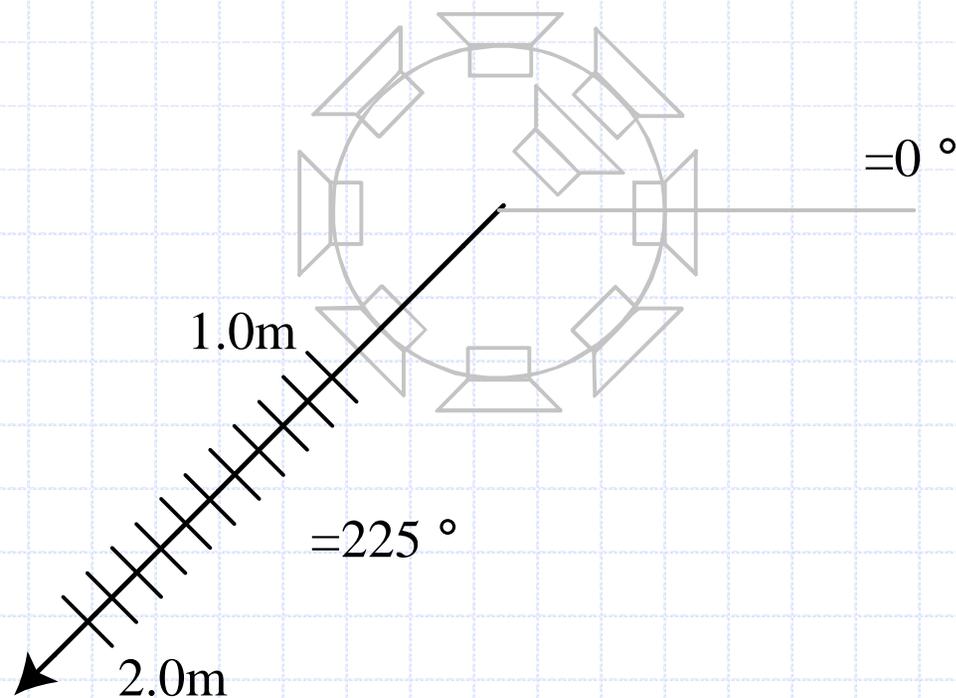
聴取位置での再現精度



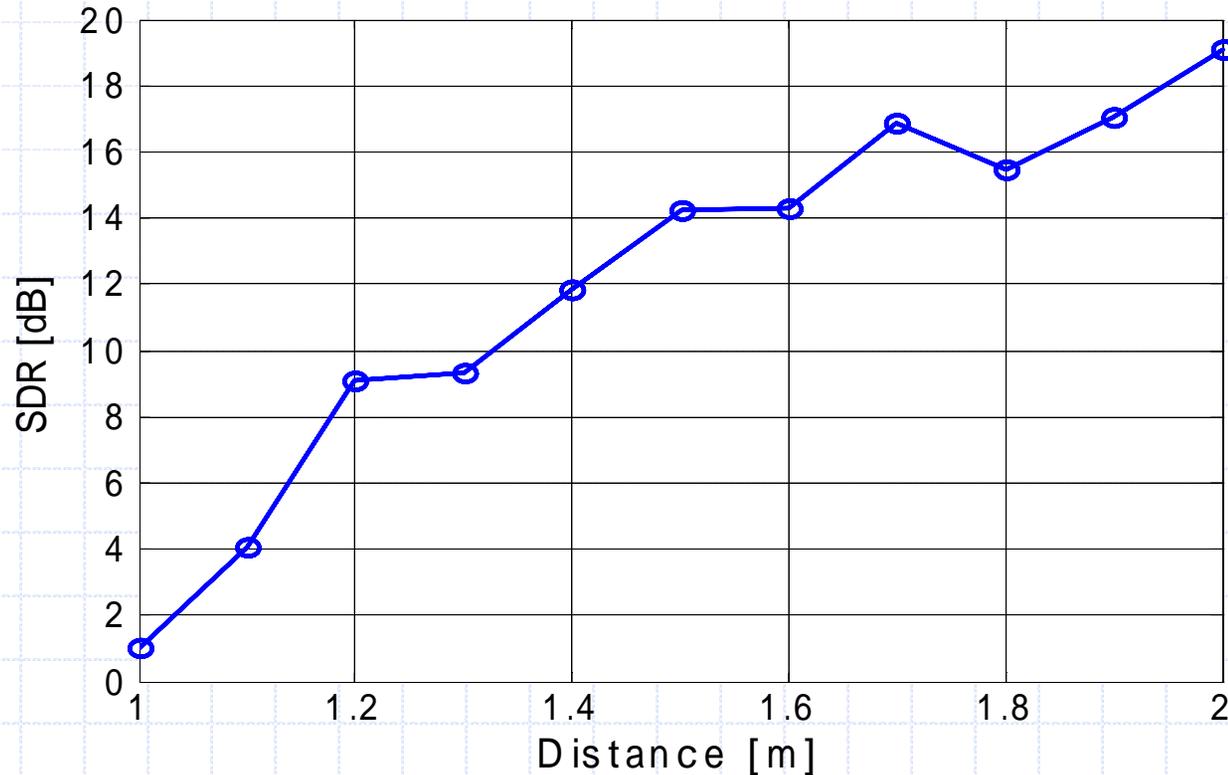
- 仮想音源に近い方位では8dBから15dB程度
- 仮想音源から遠い方位では距離による差が大きい

距離による再現精度

- 仮想音源数
 - ◆ N=36に固定
- 方位角
 - ◆ $=225^\circ$
- 中心からの距離
 - ◆ 1.0mから2.0m
 - ◆ 0.1m刻み



距離による再現精度 =225°



- 距離が離れるとSDRは大きく上昇

バイノーラル信号の評価

◆ 主観評価実験

- 被験者 5名、1名4試行
- 生成した信号とHRTFのみを畳み込んだ信号の比較
 - ◆ 実音源とHRTFを用いた信号と同じ方向に知覚されるか

音圧レベル	61.2dB(A)
サンプリング周波数	32kHz
暗騒音レベル	18.5dB(A)
方位角	$\theta = 0^\circ, 45^\circ$
距離	1.0m、1.6m
仮想音源数	12、18、24、36、48

仮想音源数と方向定位1

■ 仮想音源数

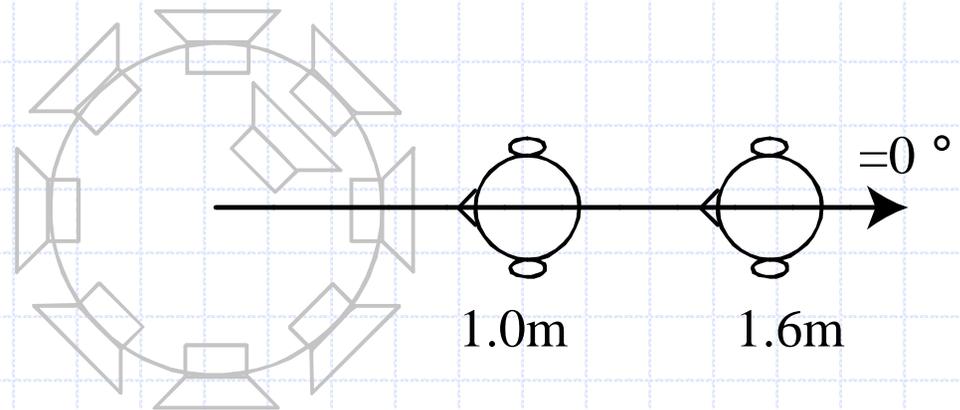
- ◆ N=12
- ◆ N=18
- ◆ N=24
- ◆ N=36
- ◆ N=48

■ 方位角

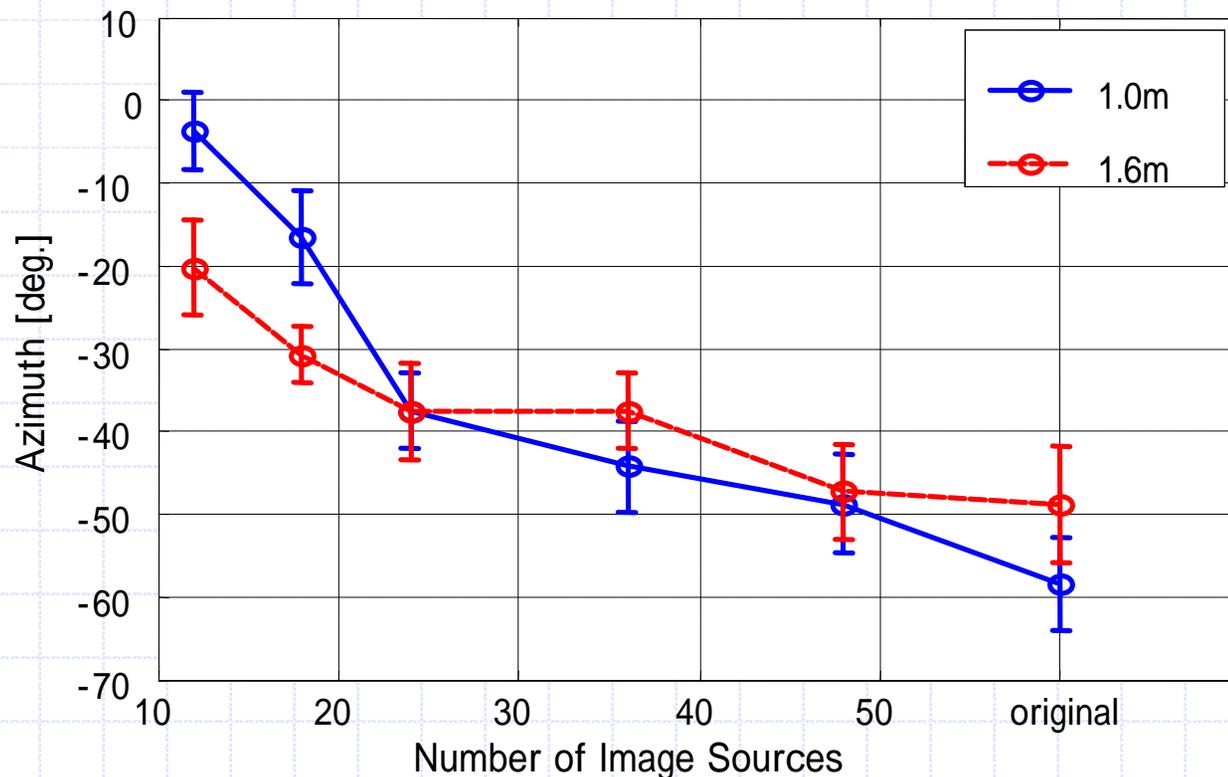
- ◆ $=0^\circ$

■ 中心からの距離

- ◆ 1.0m、1.6m



=0° 聴取者の右手前に音源を配置



- 仮想音源が増えると正しい方向に近づく
- 仮想音源から離れると正しい方向に近づく

仮想音源数と方向定位2

■ 仮想音源数

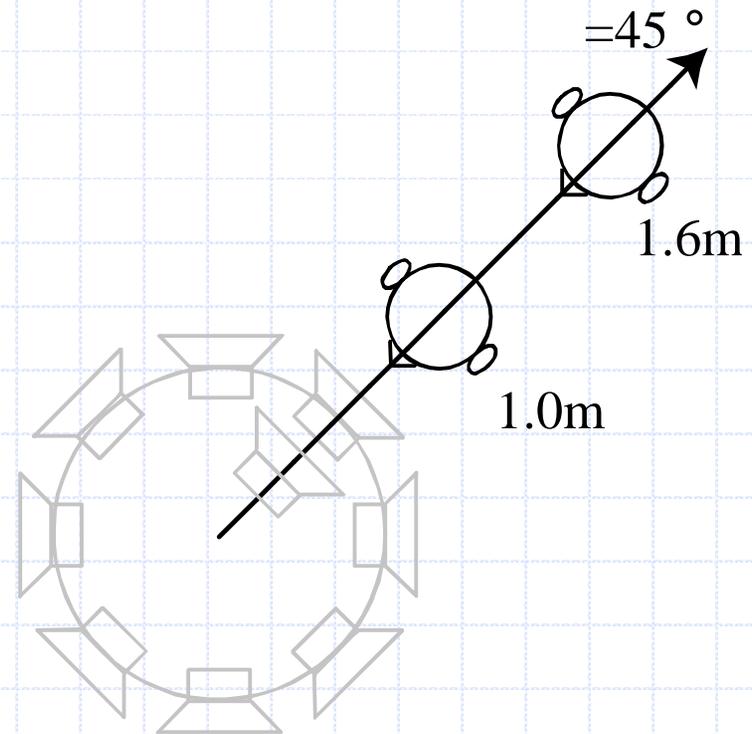
- ◆ N=12
- ◆ N=18
- ◆ N=24
- ◆ N=36
- ◆ N=48

■ 方位角

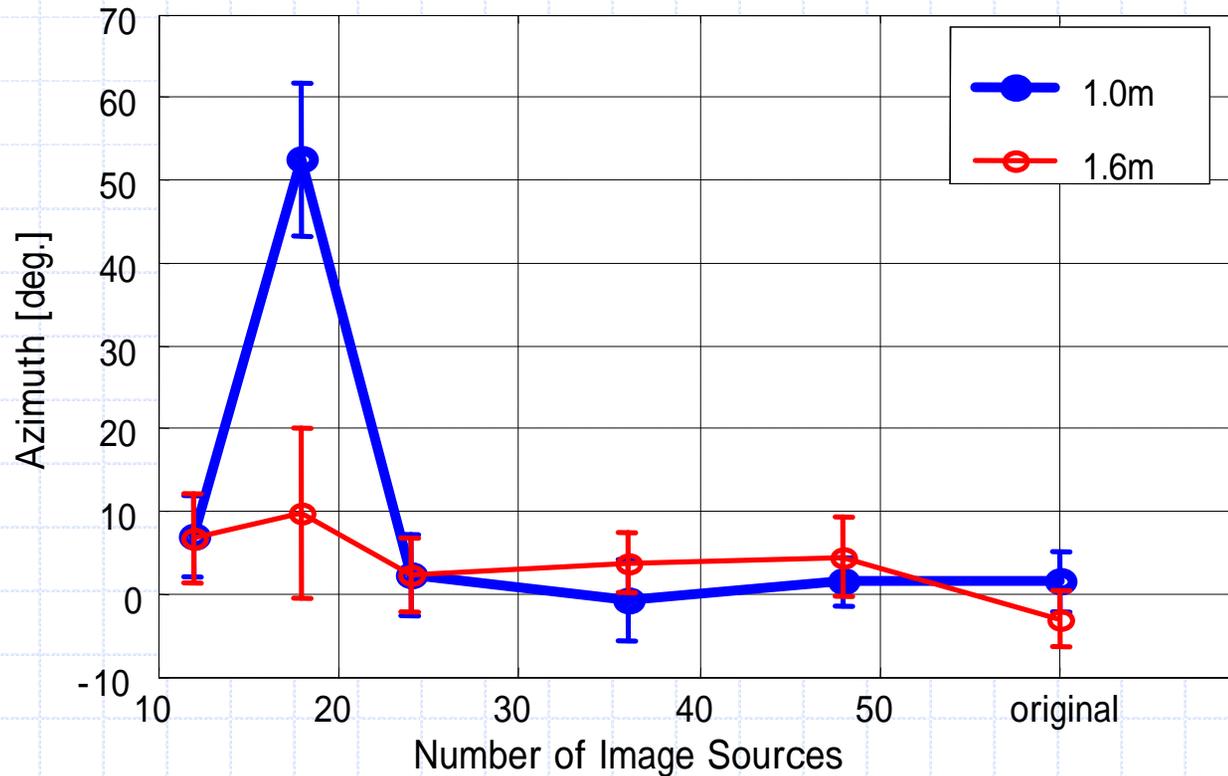
- ◆ $=45^\circ$

■ 中心からの距離

- ◆ 1.0m、1.6m



=45° 聴取者正面に音源を配置



- 仮想音源が少なく、仮想音源に近いときに正しく知覚できない領域が存在する

原因の検討

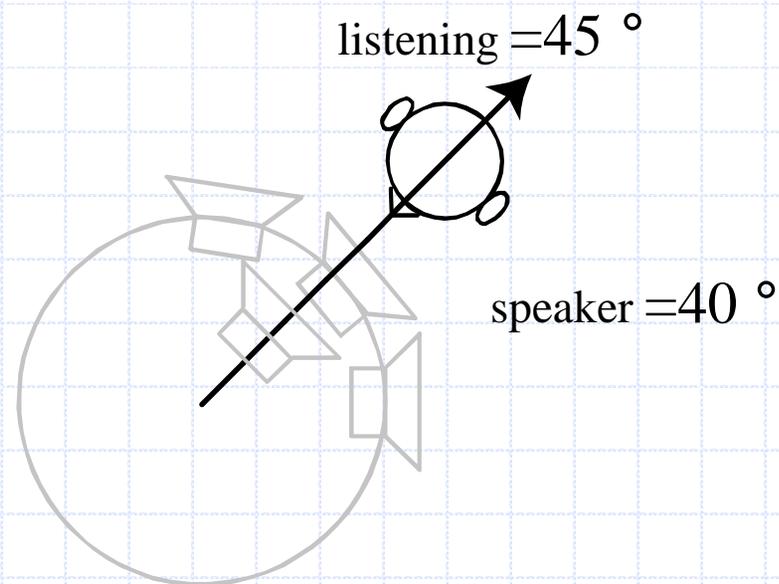
- 仮想音源の位置
 - ◆ 聴取者正面に対して非対称
- 仮想音源の数
 - ◆ 見かけ上、近距離時に減少



- 最も近い仮想音源位置に音像が知覚される



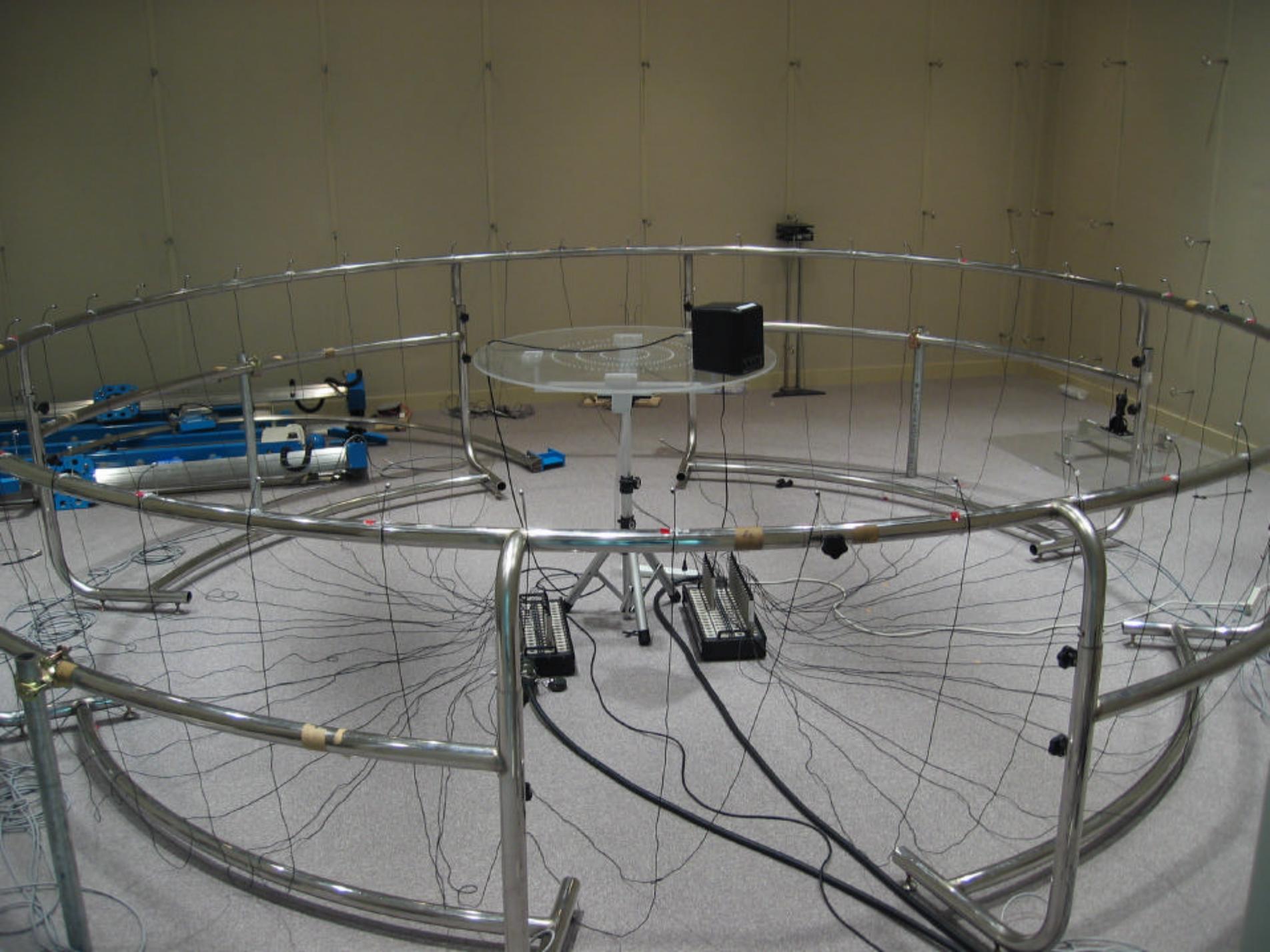
- 改善するには
 - ◆ 仮想音源数を増やす
 - ◆ 仮想音源から離れる



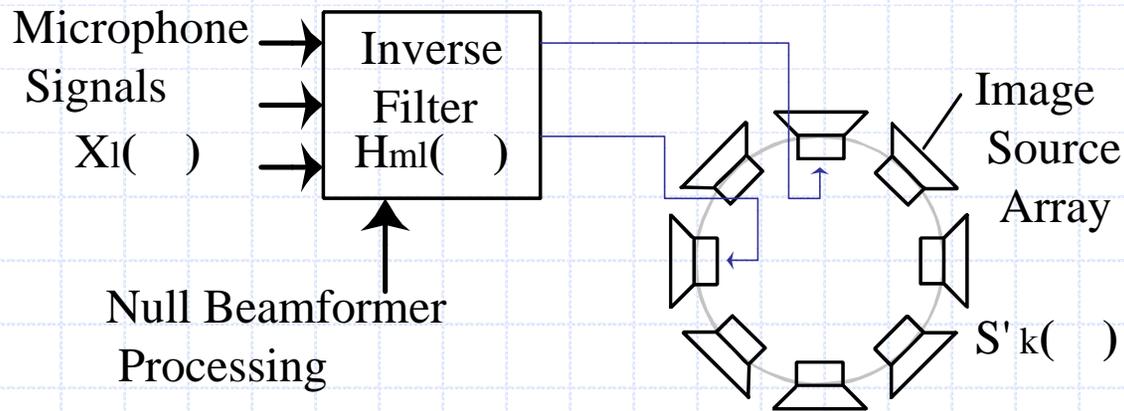
まとめ

- ◆ 自由聴点音場再生システムの構築
 - 音源を囲むようにマイクロホンアレーを配置
 - マイクロホンアレーの内側の音場再現
 - ◆ 自由視点テレビシステムに聴覚情報を与える
 - 内部に仮想音源を仮定し、推定 ➡ 音場再現
- 仮想音源の推定精度はSDRで4dBから5dB
 - ◆ マイクロホンは仮想音源の4倍必要
- 仮想音源を増やすことでより高い音場再現精度
- 主観実験により提案手法の性能を検討
- 定位方向の誤差は 10° 以内





仮想音源信号の推定



◆ 逆伝達特性 H の算出

- 伝達関数 G の逆特性

◆ 仮想音源信号 S' の推定

- マイクロホン信号 X と逆伝達特性 H より算出

逆伝達関数の算出

$$X_l(\omega) = \sum_{k=1}^N G_{kl}(\omega) S_k(\omega) \quad \left| \quad S'_m(\omega) = \sum_{l=1}^M H_{lm}(\omega) X_l(\omega) \right|$$

$$\sum_{l=1}^M H_{lm}(\omega) G_{kl}(\omega) = \begin{cases} S'_k(\omega)/S_k(\omega) & (k = m) \\ 0 & (k \neq m) \end{cases}$$

$$S'_k(\omega)/S_k(\omega) = e^{-j\omega\tau_0}$$

$$\mathbf{H}(\omega)\mathbf{G}(\omega) = \mathbf{D}(\omega)$$

$$\mathbf{H}(\omega) = \mathbf{D}(\omega)\mathbf{G}^+(\omega)$$

今後の課題

- ◆ 複数音源への対応
- ◆ 自由視点テレビシステムとの連動

仮想音源数による再現精度

■ 仮想音源数

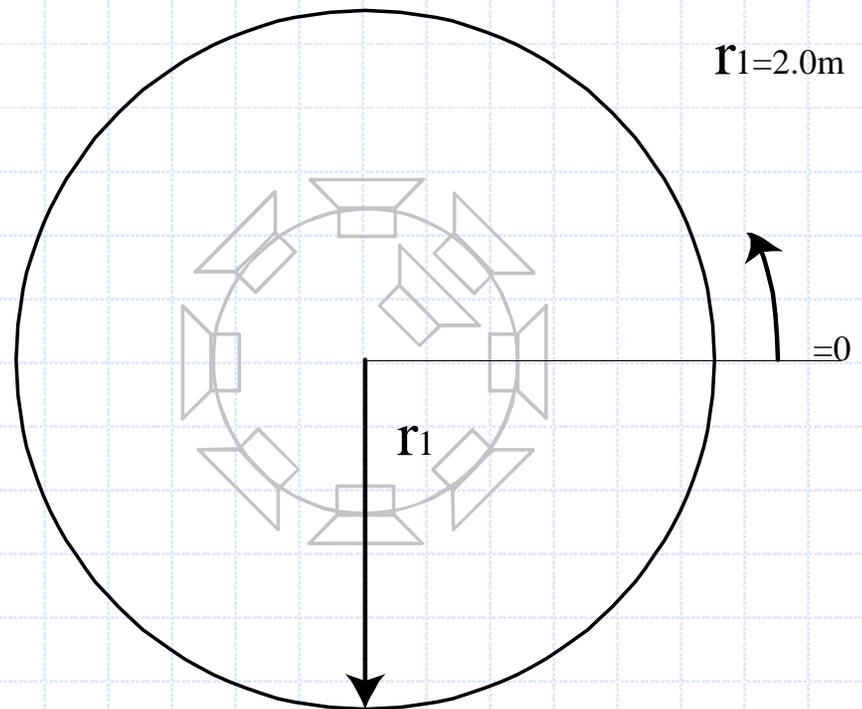
- ◆ $N=12$
- ◆ $N=18$
- ◆ $N=24$
- ◆ $N=36$
- ◆ $N=48$

■ 円周上で再現

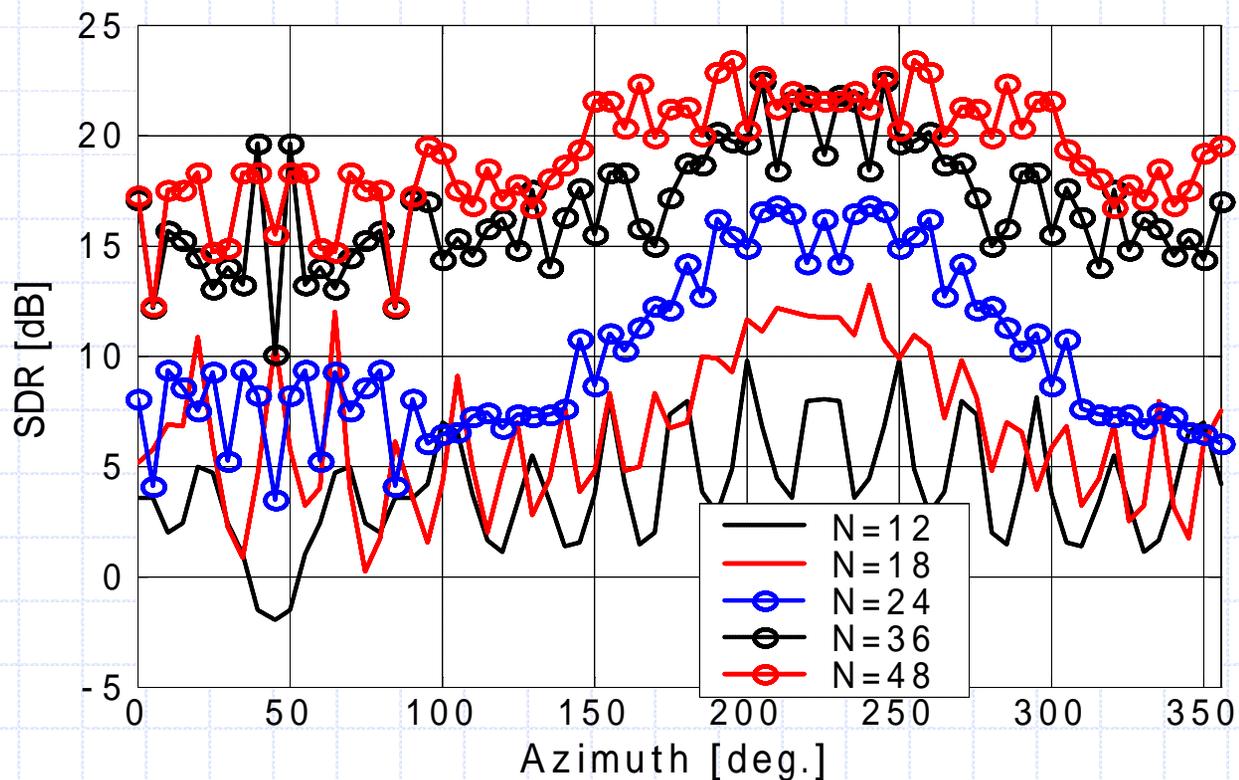
- ◆ 方位各 5° 刻み

■ 半径

- ◆ 2.0m



仮想音源数別の再現精度



- 仮想音源の増加はSDRの上昇に関係

距離による再現精度1

■ 仮想音源数

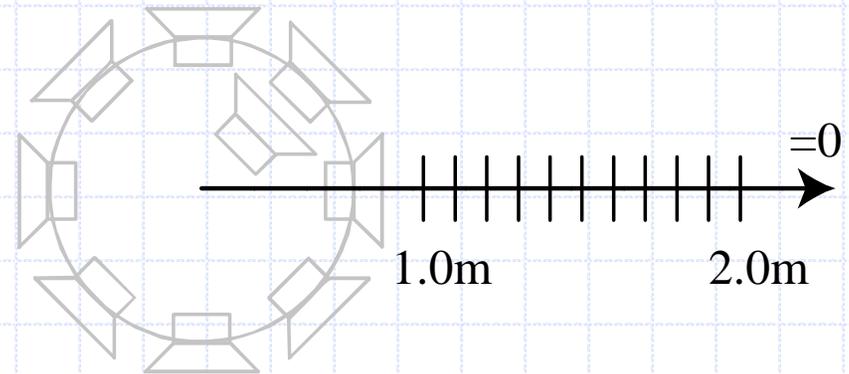
- ◆ N=12
- ◆ N=18
- ◆ N=24
- ◆ N=36
- ◆ N=48

■ 方位角

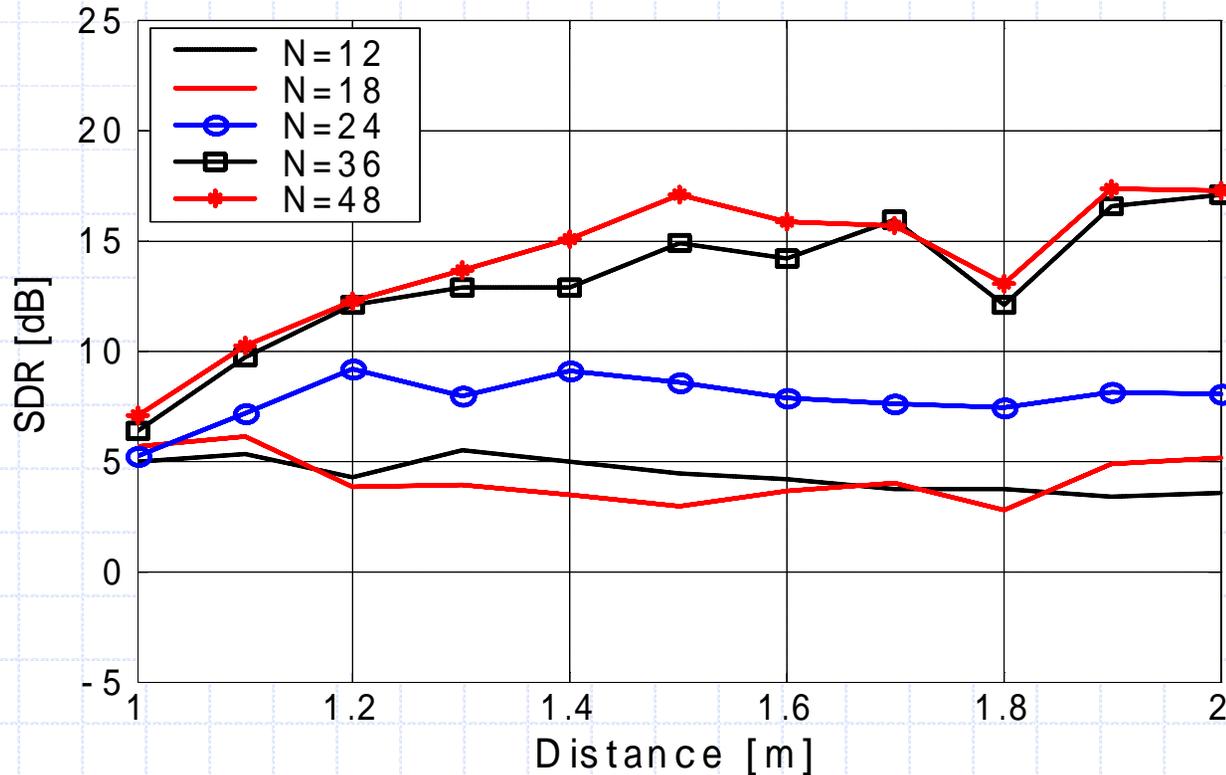
- ◆ $=0^\circ$

■ 中心からの距離

- ◆ 1.0mから2.0m
- ◆ 0.1m刻み



$\approx 0^\circ$

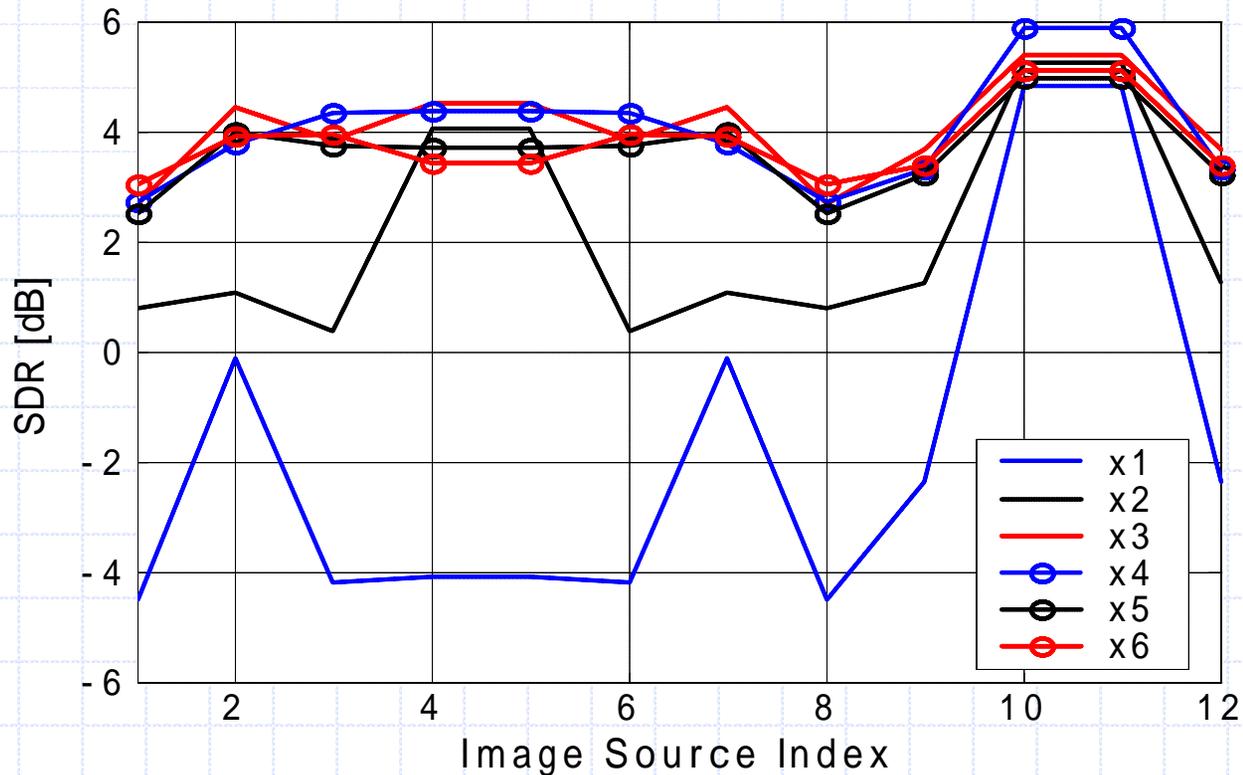


- 仮想音源が少ない
- 仮想音源が多い

距離による変化なし

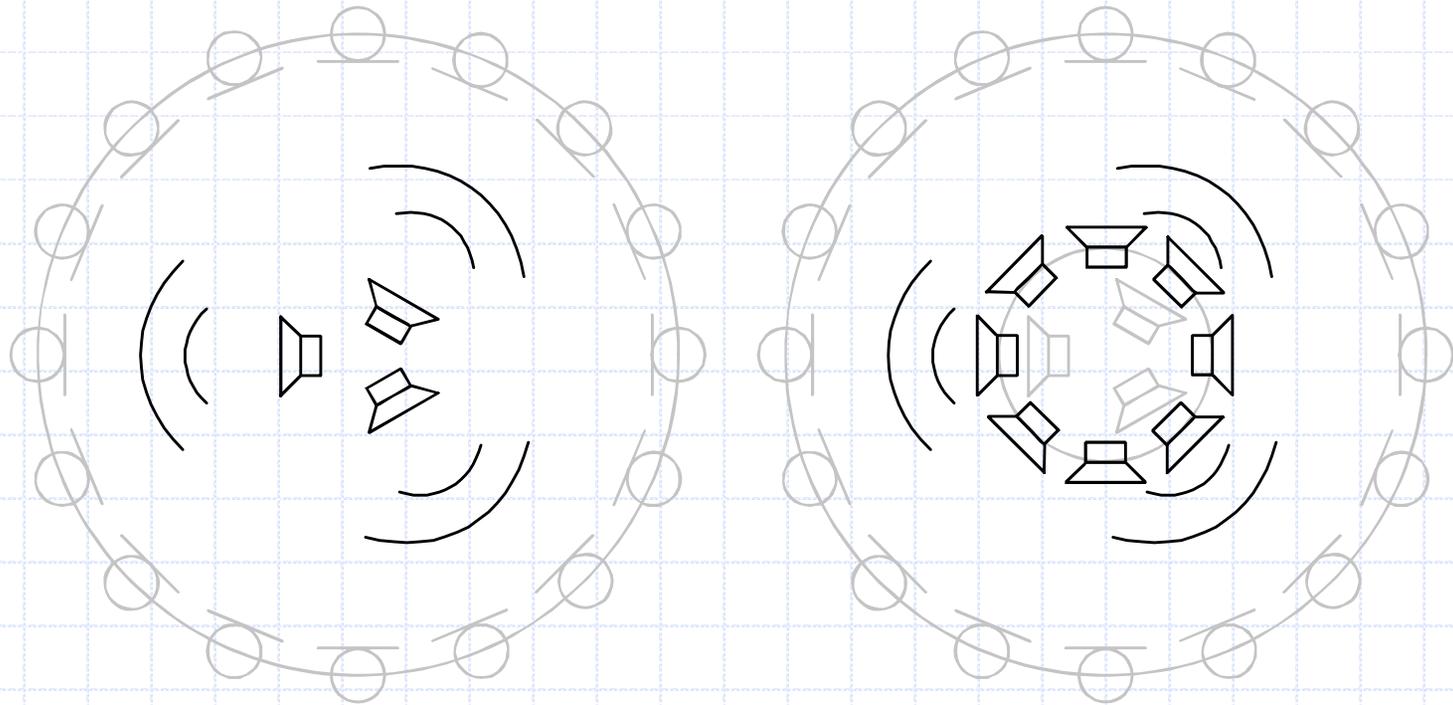
距離によってSDRが変化

仮想音源信号の推定とマイク数の比



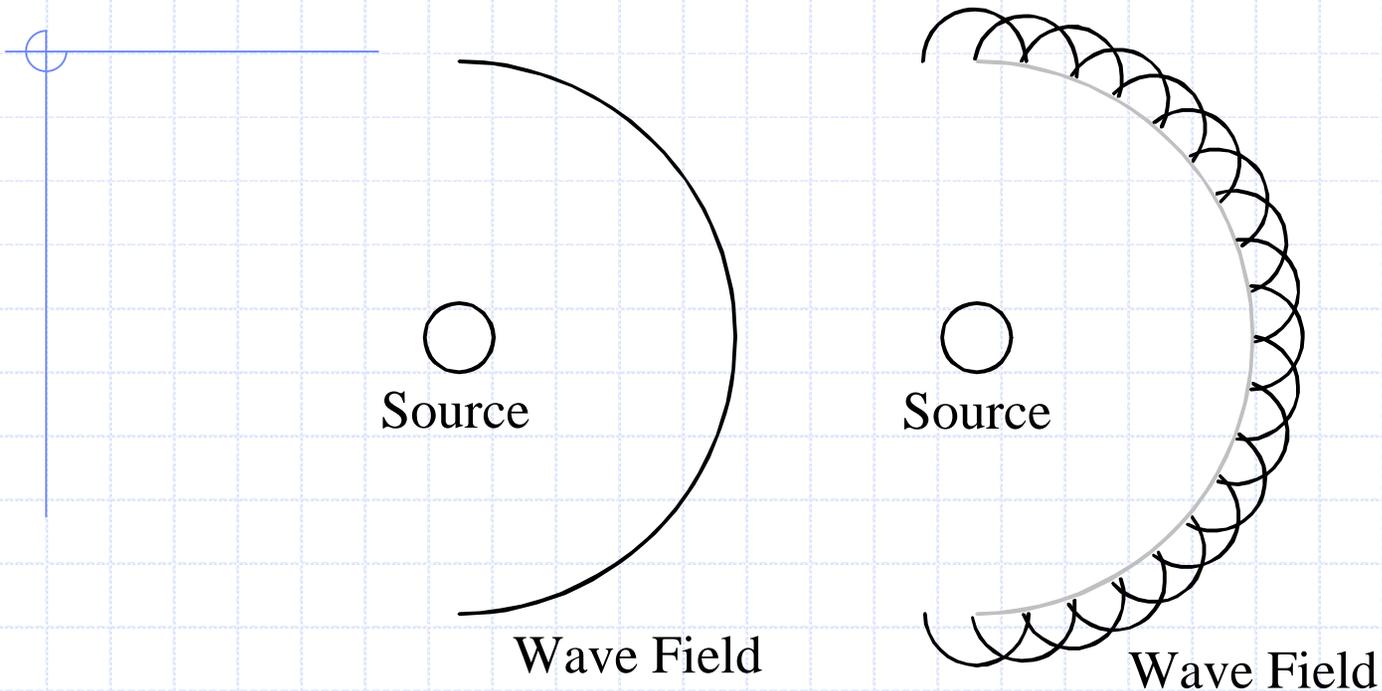
- マイクロホンは仮想音源の4倍必要

2次音源による再現



- カメラアレーと同一位置に配置したマイクロホンアレー信号を用いて2次音源信号を推定
 - ◆ 仮想音源と呼称

Huygensの原理



- 波面が進むとき、ある瞬間における各点は新しい波面となる
- これらの素元波を包むように接する面が次の波面となる

研究の背景

◆ 音場再現手法

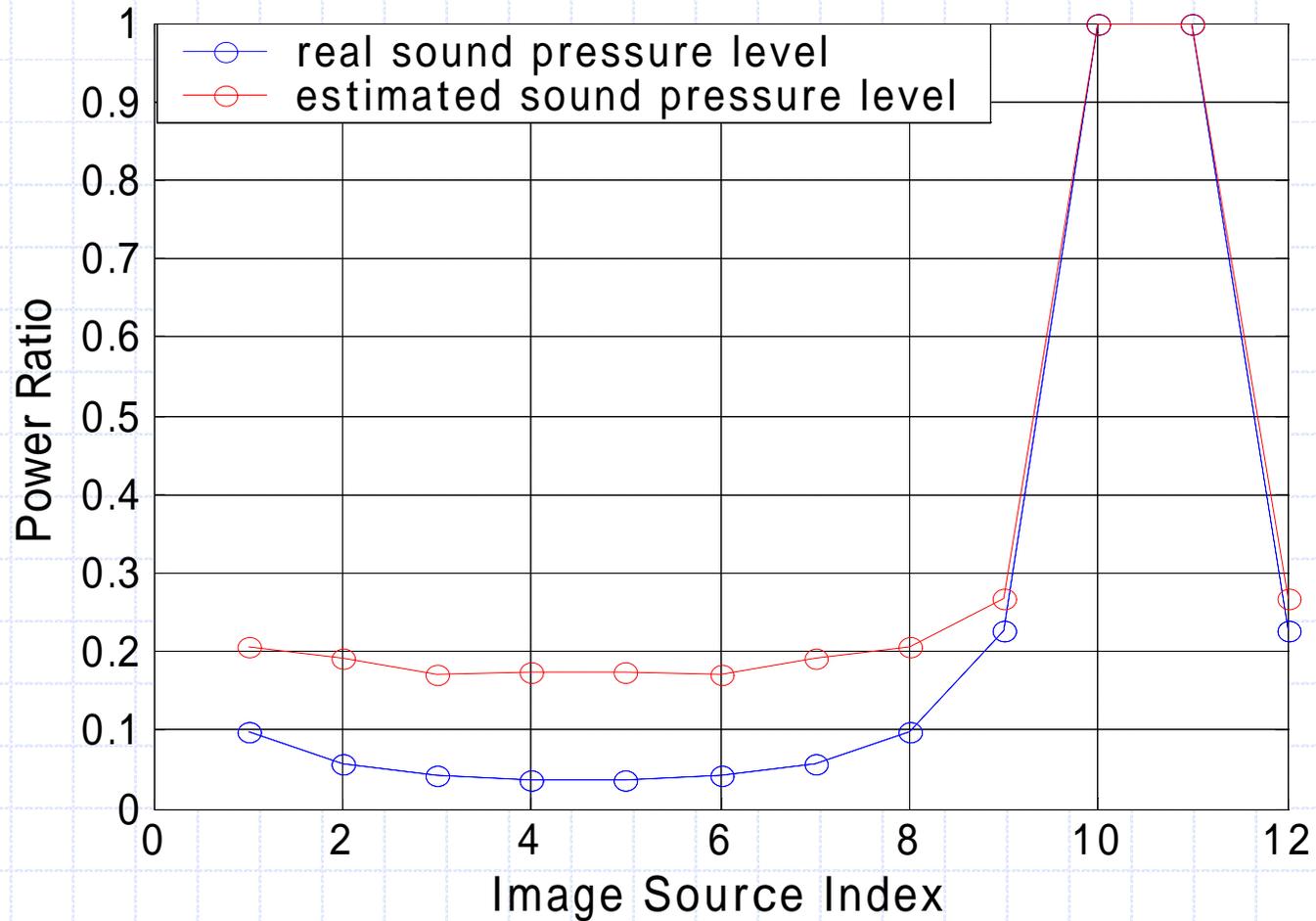
- バイノーラル方式
 - ◆ ヘッドホンによる提示
- トランスオーラル方式
 - ◆ スピーカによる提示
- 波面合成
 - ◆ 2次音源による実音源の置換
- 境界音場制御
 - ◆ 境界上の音圧と音圧勾配を等しくする

両耳位置の制御

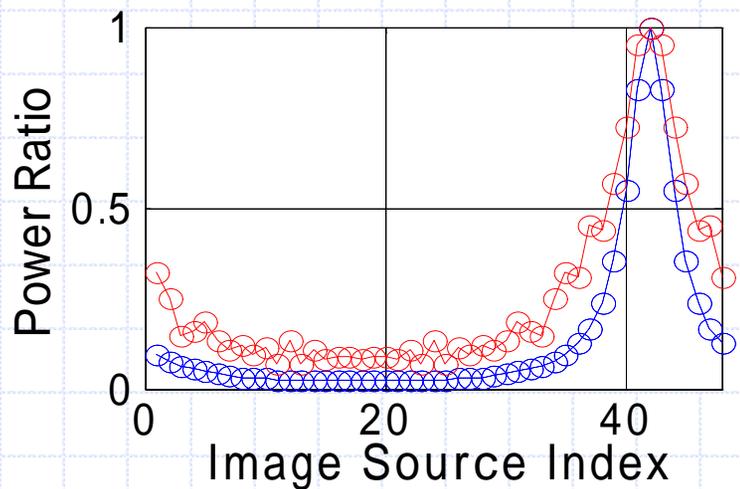
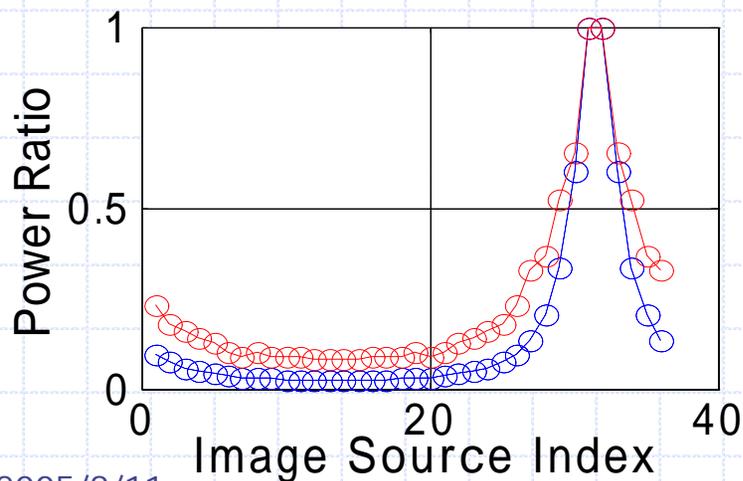
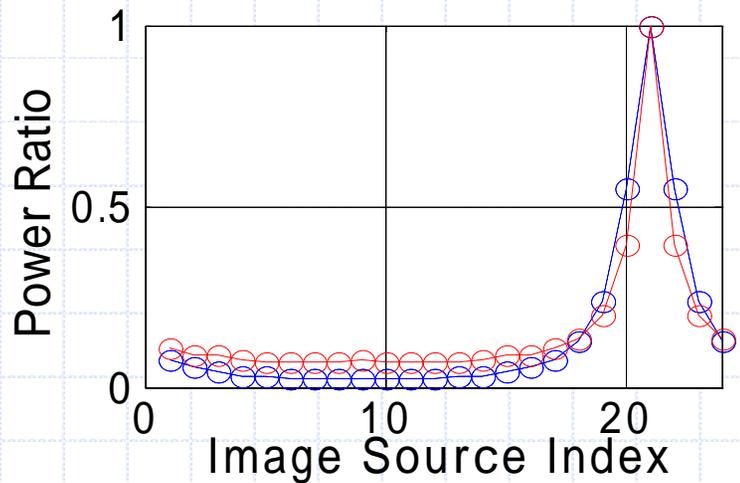
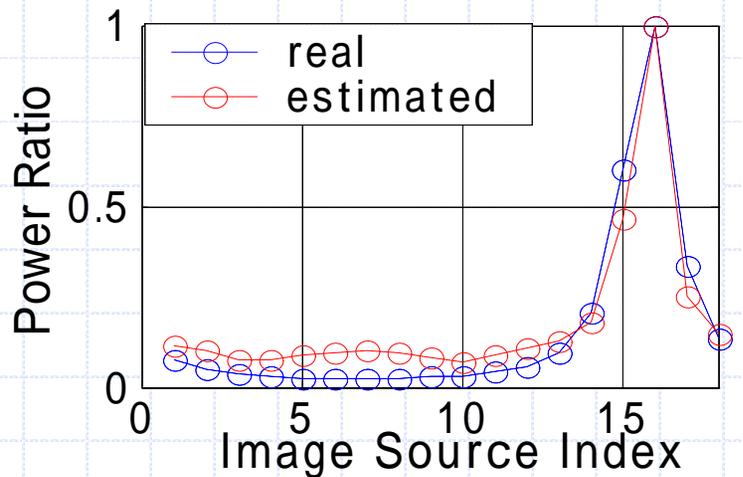
領域の制御

◆ 音場再現を用いた立体音場システムの構築

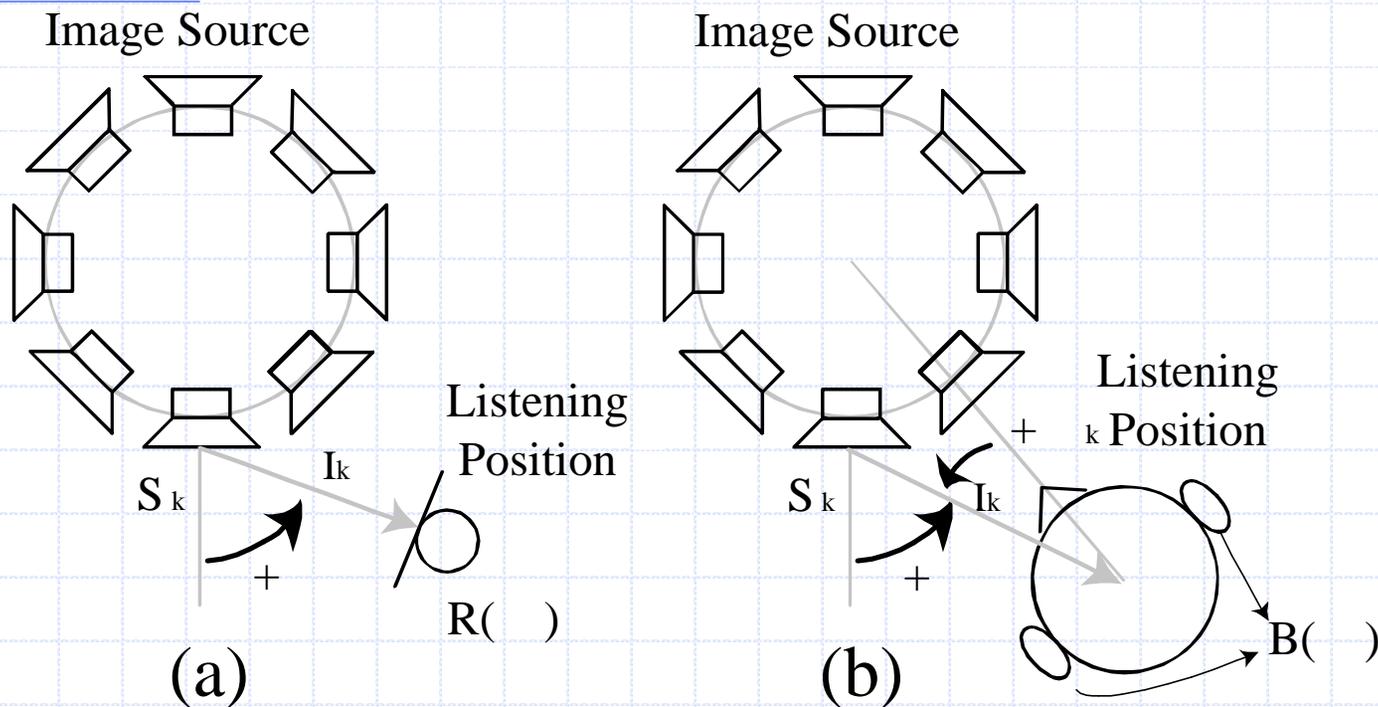
仮想音源信号のパワー比



仮想音源信号のパワー比



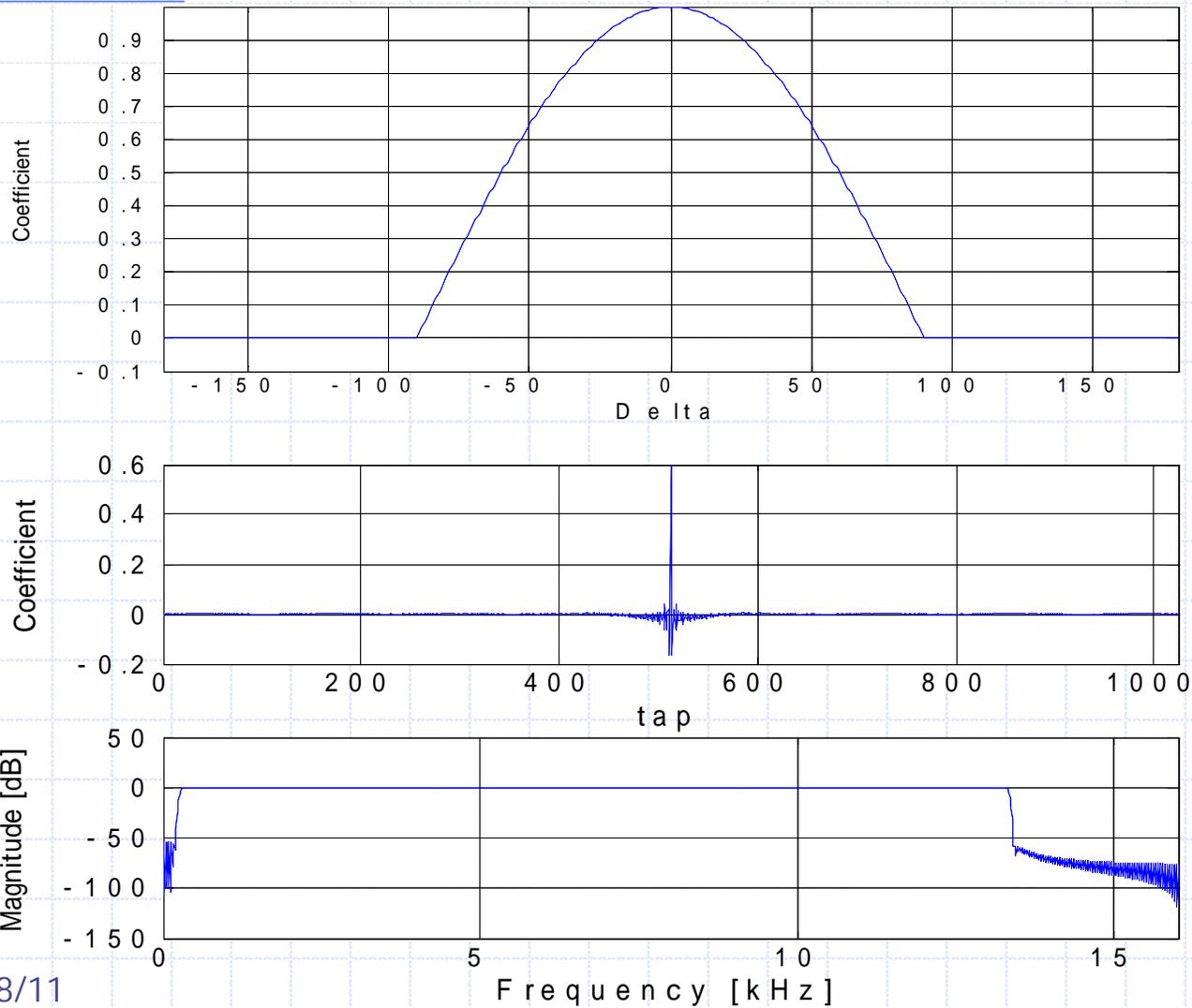
聴取位置での信号生成



$$R(\omega) = \sum_{k=1}^N c_k(\Delta) I_k(\omega) S_k(\omega)$$

$$B(\omega) = \sum_{k=1}^N c_k(\Delta) D_k(\phi_k, \omega) I_k(\omega) S_k(\omega)$$

指向特性とFIRバンドパスフィルタ

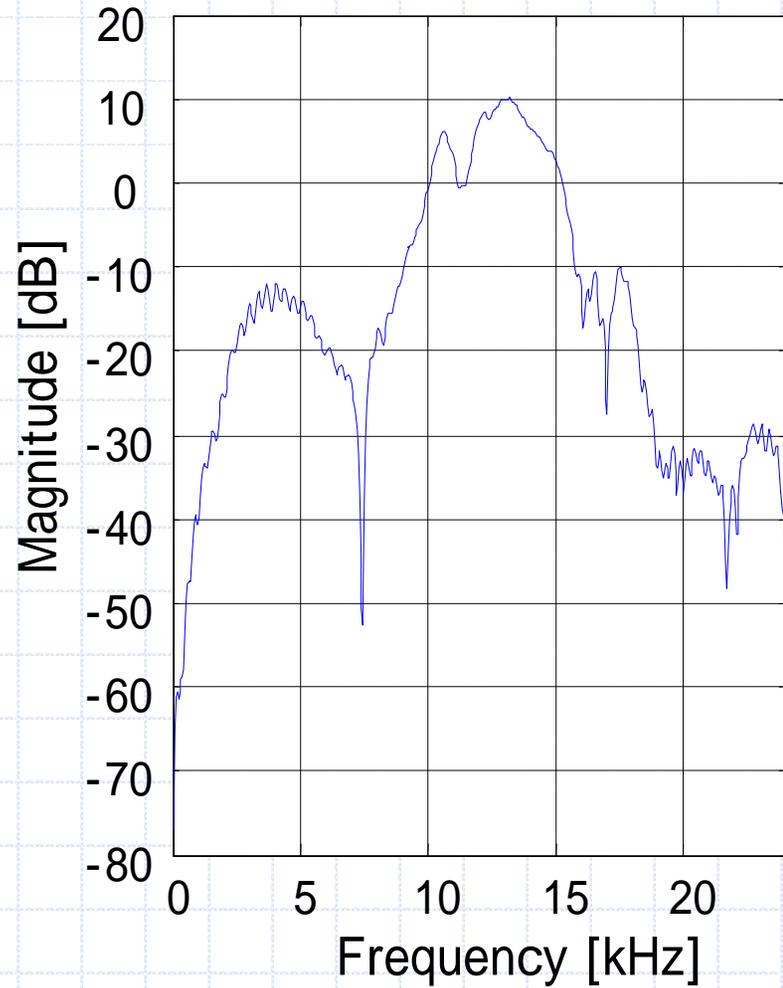
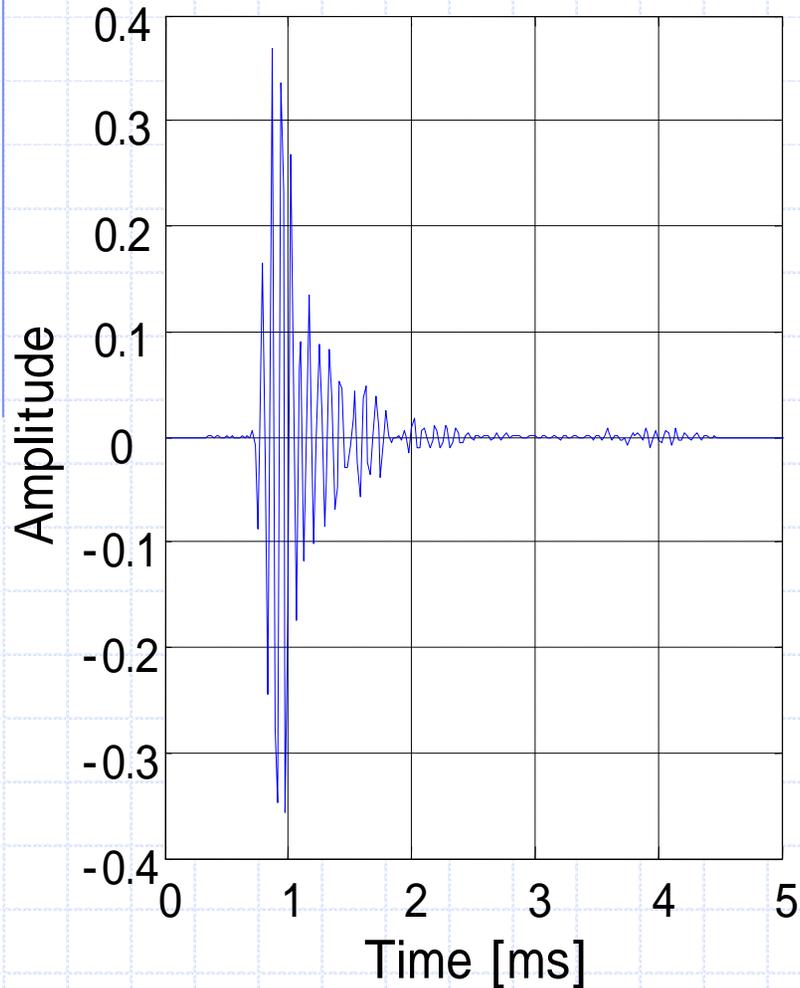


近距離HRTFの測定

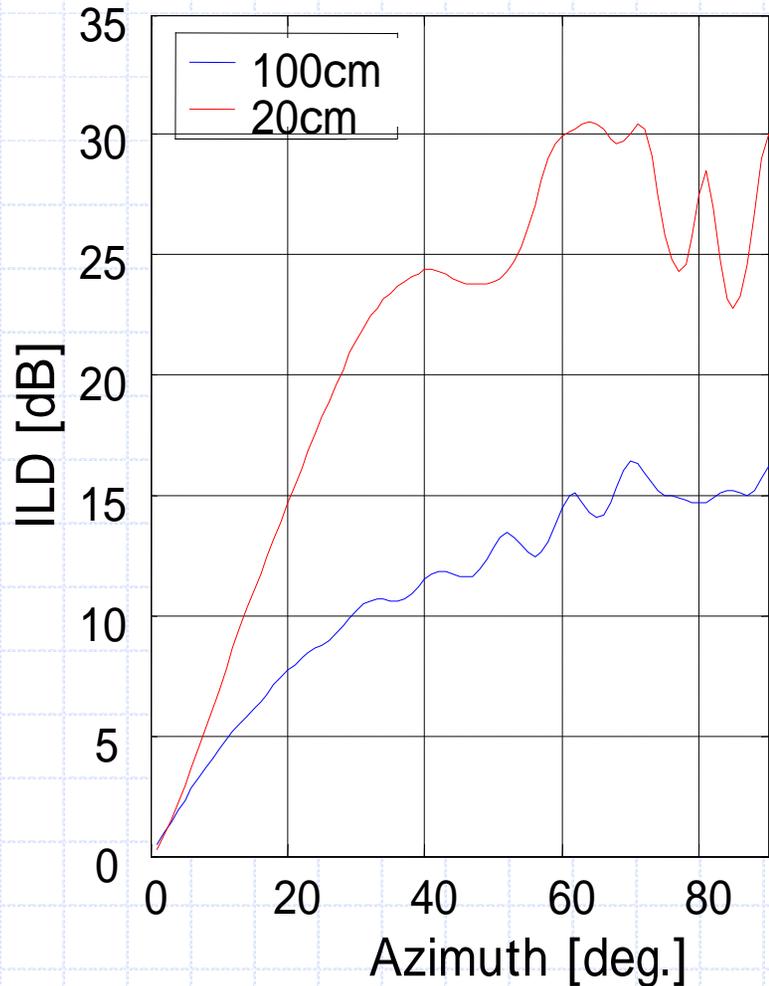
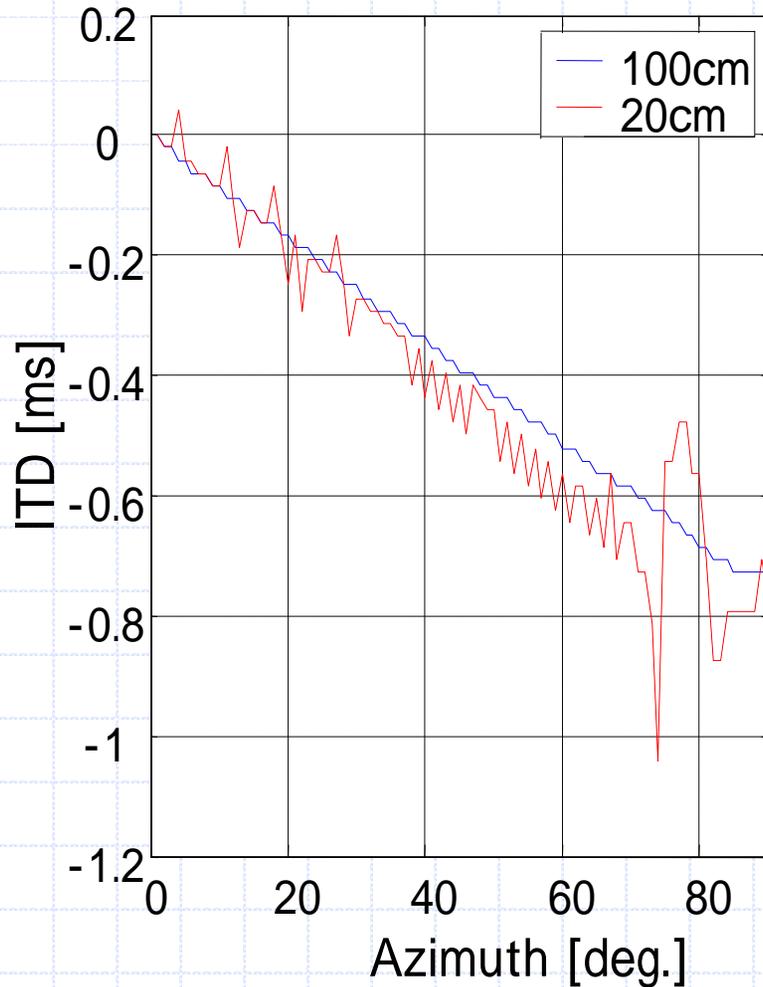


音圧レベル	69.0dB(A) (1m)
サンプリング周波数	48kHz
暗騒音レベル	13.8dB(A)

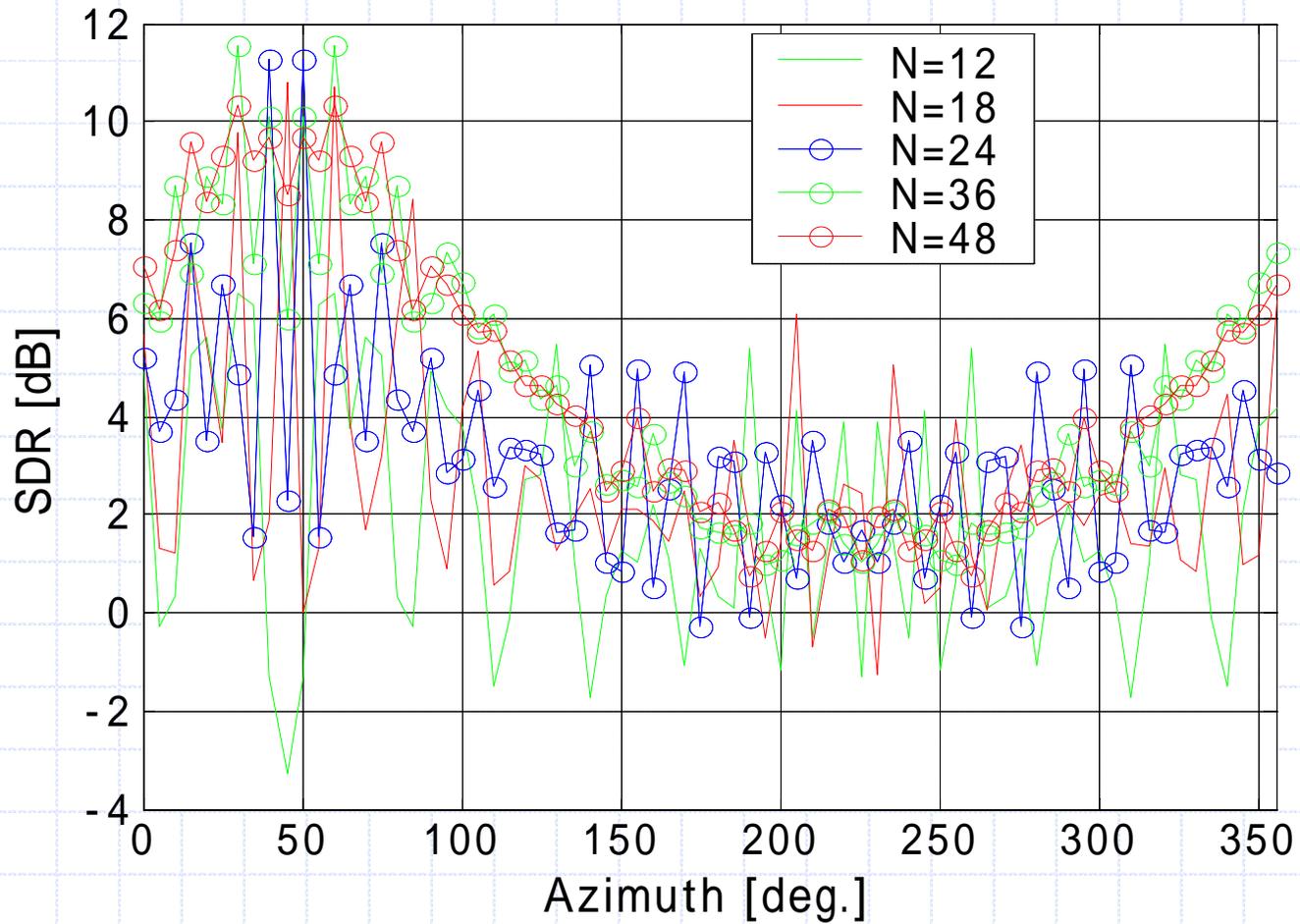
HRTFの波形



ITDとILDの比較



R=1m



Theta=45

