

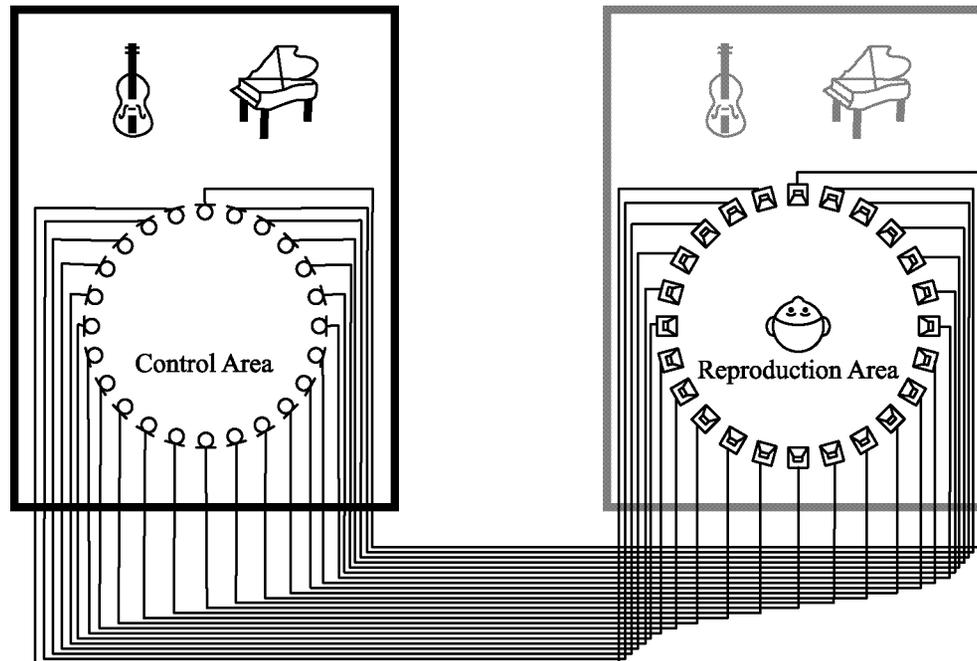
逆フィルタを用いた 音場空間符号化方式の主観評価

木村敏幸, 笥一彦(名大・人間情報/CIAIR),
武田一哉, 板倉文忠(名大・工/CIAIR)

1 .はじめに

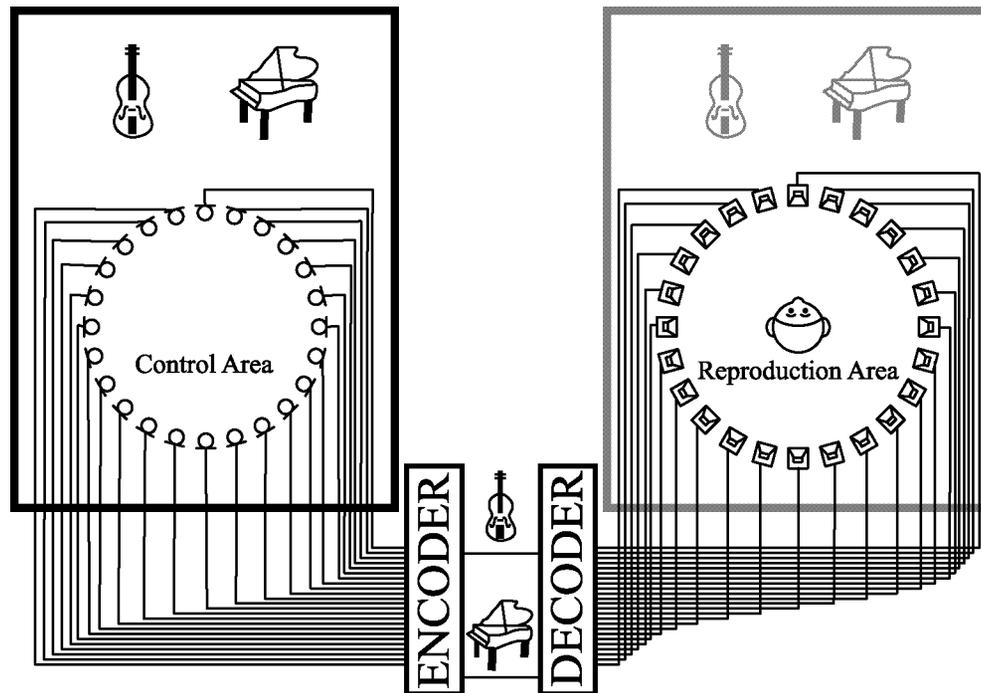
研究背景

- スピーカーアレーを用いた音場再現技術
 - 再現領域で制御領域の音響情報を再現
 - 伝送量がチャンネル数に比例
 - 従来の符号化方式では伝送量の削減が不十分



本報告の目的

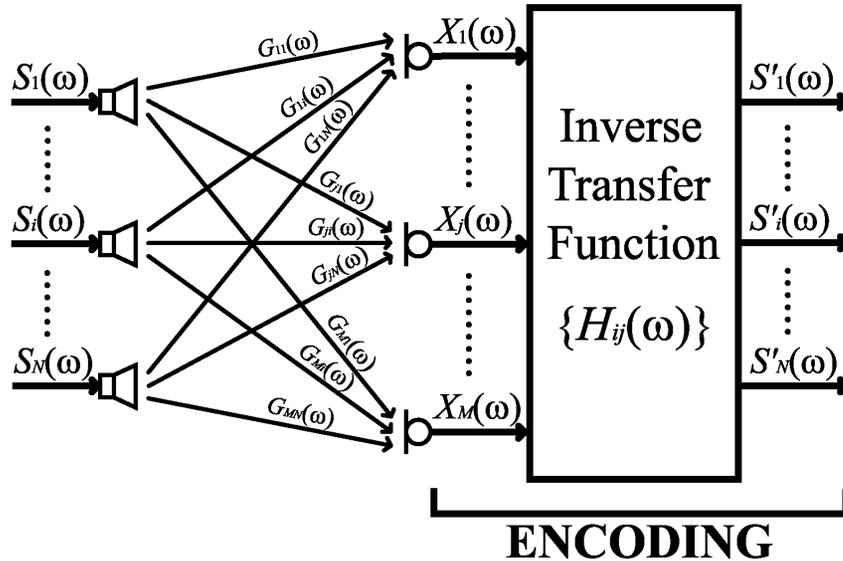
- 今までに
 - 音源抽出に基づく空間符号化方式を提案
- 今回
 - 主観評価によって符号化性能を知覚的に検討



2. アルゴリズム

符号化

- 逆伝達関数を畳み込んで、音源を抽出



$S_i(\omega)$: 音源信号

$X_j(\omega)$: チャンネル信号

$S'_i(\omega)$: 抽出した音源信号

$G_{ji}(\omega)$: $S_i(\omega)$ から $X_j(\omega)$ への室内伝達関数

$H_{ij}(\omega)$: $X_j(\omega)$ から $S'_i(\omega)$ への逆伝達関数

$$\begin{pmatrix} H_{11}(\omega) & \cdots & H_{N1}(\omega) \\ \vdots & \ddots & \vdots \\ H_{1M}(\omega) & \cdots & H_{NM}(\omega) \end{pmatrix} = \begin{pmatrix} G_{11}(\omega) & \cdots & G_{M1}(\omega) \\ \vdots & \ddots & \vdots \\ G_{1N}(\omega) & \cdots & G_{MN}(\omega) \end{pmatrix}^+ \begin{pmatrix} S'_1(\omega) & \cdots & 0 \\ S_1(\omega) & \ddots & \vdots \\ 0 & \cdots & \frac{S'_N(\omega)}{S_N(\omega)} \end{pmatrix} \Leftrightarrow \mathbf{H}(\omega) = \mathbf{G}^+(\omega)\mathbf{D}(\omega)$$

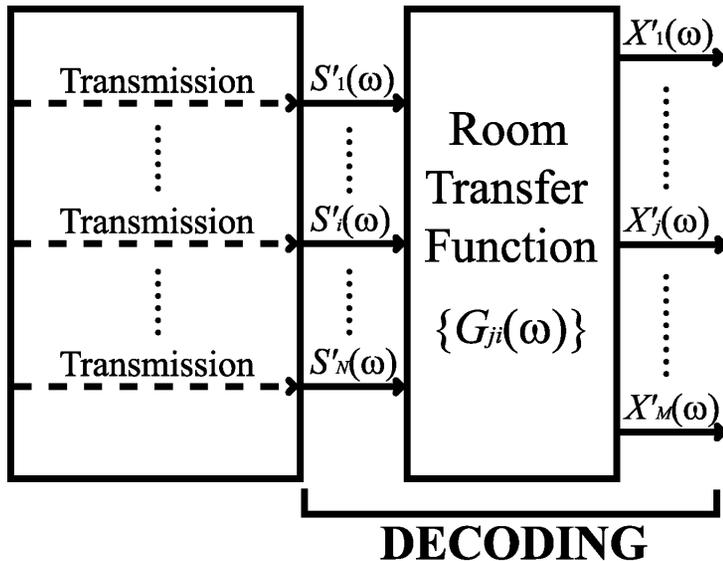
$\mathbf{G}(\omega)$: 室内伝達関数行列

$\mathbf{G}^+(\omega)$: $\mathbf{G}(\omega)$ の Moore-Penrose 擬似逆行列

$\mathbf{H}(\omega)$: 逆伝達関数行列

復号化

- 室内伝達関数を畳み込んで,チャンネルを復元
 - 送信側と受信側で室内伝達関数の情報は共有



$S'_i(\omega)$: 抽出した音源信号

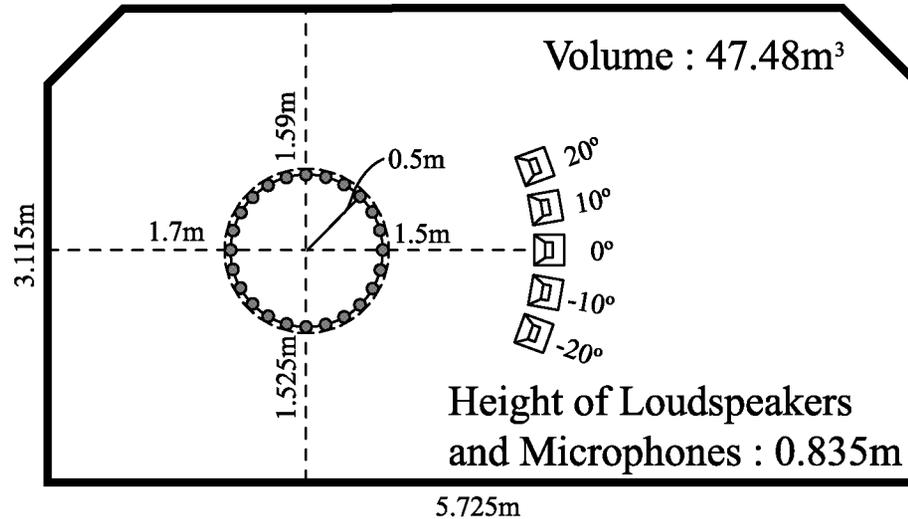
$X'_j(\omega)$: 復元したチャンネル信号

$G_{ji}(\omega)$: $S_i(\omega)$ から $X_j(\omega)$ への室内伝達関数

伝送量 : M個のチャンネル N個の音源
($M \gg N$ のとき, 特に有効)

3 .符号化実験

室内インパルス応答の測定



残響時間	150ms	300ms
室温	19.5°C	19.2°C
暗騒音レベル	20.0dB(A)	19.4dB(A)
音圧レベル*1	90.0dB(A)	91.6dB(A)
FIRフィルタ次数	7200	14400

標本化周波数...48kHz

手法...TSP法

TSPの長さ...65536点

同期加算...16

*1 スピーカーから1m

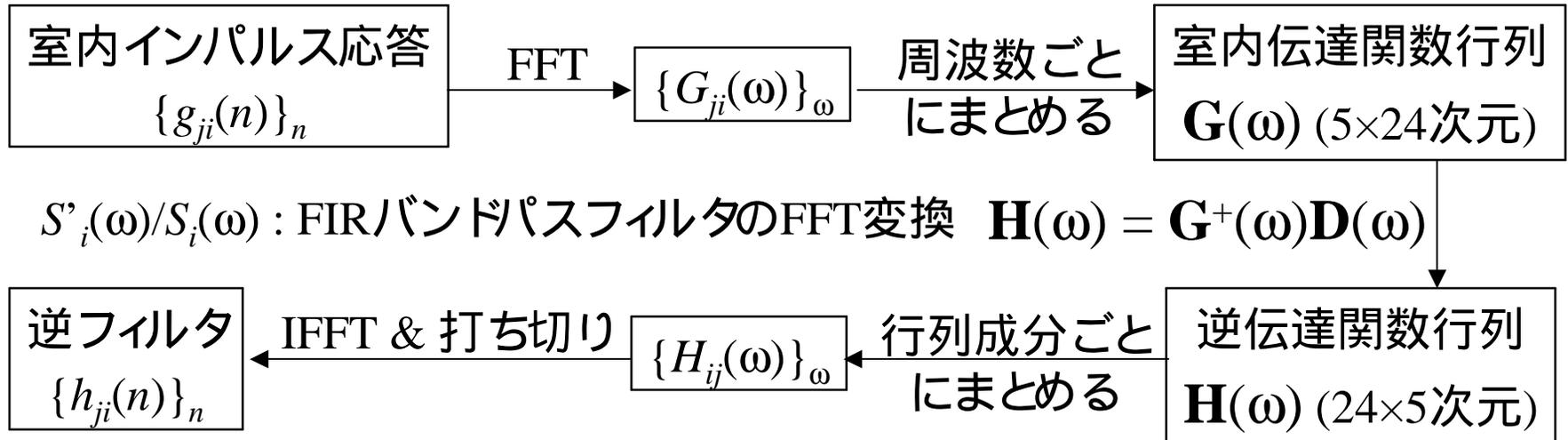
マルチチャンネル信号の作成

音源信号	女性音声	木琴	木管五重奏
長さ	5秒		
標本化周波数	12kHz		
帯域	50-5000Hz		
音源の数	1	1	5
音源の方位角	0°	0°	0°, ±10°, ±20°
残響時間	0.6sec	1.2sec	1.2sec

- 残響時間

– 150ms & 300ms 0.6sec & 1.2sec

逆フィルタの算出



残響時間	0.6sec	1.2sec
FFT points	32768	65536
標本化周波数	12kHz	
BPFの帯域	50-5000Hz	
BPFの遅延	10ms, 20ms, 40ms	
逆フィルタ長	14400	28800

BPFの遅延

逆フィルタの因果律を満たすために必要

符号化遅延に相当

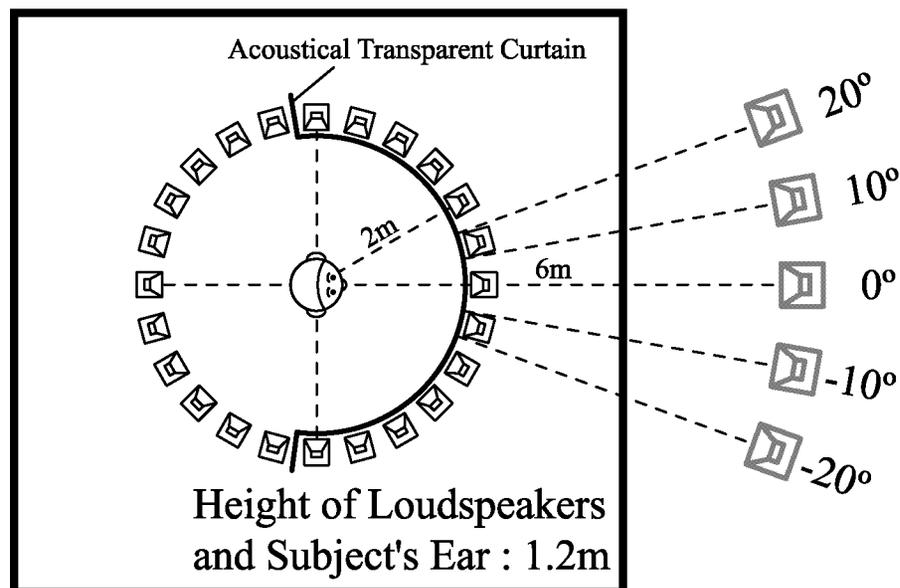
符号化遅延が短いほどより実用的

4 .主観評価実験

主観評価実験 (環境)

- 低残響室 (残響時間...約80ms)
 - 暗騒音レベル...25.0dB(A)
 - 音圧レベル...受聴位置で約70dB(A)
 - 被験者...学生6名 (男性5名, 女性1名)

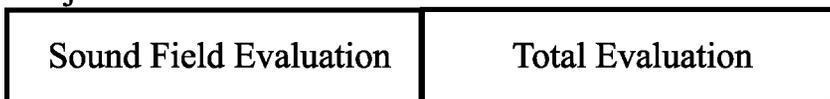
Low-Reverberation Room



主観評価実験 (構成)

- 手法...三刺激二重盲験法 (XAB法)

Subjective Assessment



Evaluation



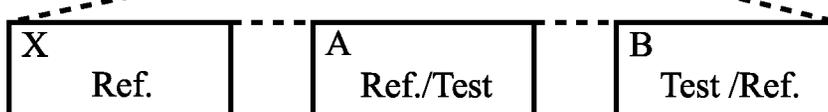
Randomize

(Speech, Xylophone, Woodwind Quintet)

Session



Trial



Break
(0.5sec)

Break
(0.5sec)

X	A	B	練習	本番
原音	原音	符号化音 (遅延10ms)	1	2
原音	符号化音 (遅延10ms)	原音	1	2
原音	原音	符号化音 (遅延20ms)	1	2
原音	符号化音 (遅延20ms)	原音	1	2
原音	原音	符号化音 (遅延40ms)	1	2
原音	符号化音 (遅延40ms)	原音	1	2

主観評価実験 (手順)

- 音場評価...部屋の知覚的な差を検討
 - X...基準の部屋で再生
 - A, B...どちらかは基準の部屋で再生
 - AとBのうち一方を5.0 ,もう一方をランク付け
- 全体評価...総合の知覚的な差を検討
 - X...原音
 - A, B...どちらかは原音と全く同じ
 - AとBのうち一方を5.0 ,もう一方をランク付け

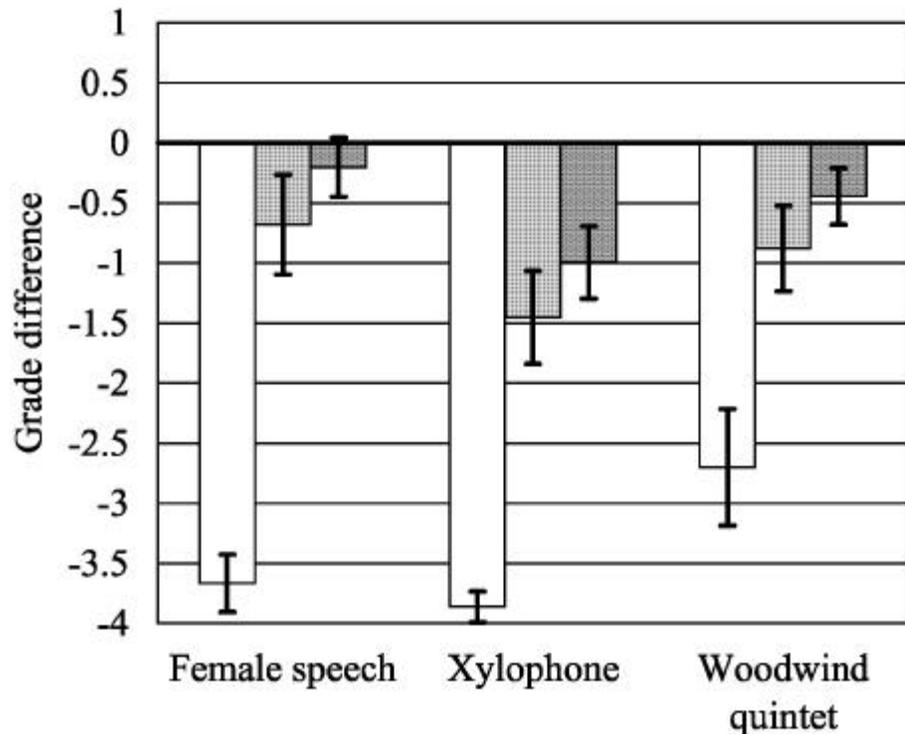
5.0	差が分からない	2.0	差が気になる
4.0	差は分かるが ,気にならない	1.0	差が非常に気になる
3.0	差がわずかに気になる		

主観評価実験 (結果)

- ランク差 = 符号化音のランク - 原音のランク

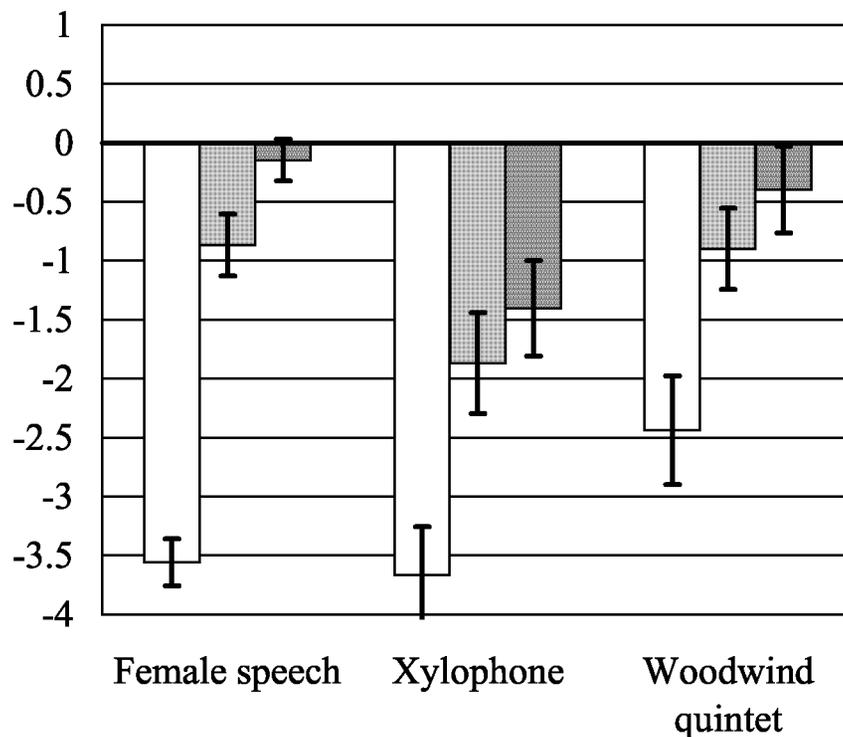
音場評価

□ Delay = 10ms ▨ Delay = 20ms ■ Delay = 40ms



全体評価

□ Delay = 10ms ▨ Delay = 20ms ■ Delay = 40ms



まとめ

- 逆フィルタを用いた空間符号化方式を提案
- 符号化による知覚歪みを主観評価によって検討
- 以下の条件であれば歪みは気にならない
 - 符号化遅延...20ms以上
 - 音源...持続する音
- 今後の予定
 - 衝撃音に対する知覚性能の改善
 - 移動音源や音源がチャンネル数より多い場合の検討