

逆フィルタを用いた 空間音響圧縮技術の拡がり感

木村敏幸, 笥一彦 (名大・人情/CIAIR)

武田一哉, 板倉文忠 (名大・工/CIAIR)

はじめに

- スピーカーアレーを用いた音場再現
 - 応用
 - 遠隔通信会議, 遠隔コンサートホール
 - 特長
 - 複数人数による受聴が容易, 拘束条件が少ない
- 遠隔での実現
 - 多数のチャンネルの伝送が必要
 - チャンネルの数が大...実現困難

研究目的

- 符号化による伝送量の削減
 - 従来の符号化技術(MPEG2 AAC, AC-3)
 - 空間上の情報圧縮においてはあまり効率的でない
- チャンネル信号
 - 音源信号に室内伝達関数の畳み込み
 - 音源信号は共通
- チャンネル信号から音源信号を抽出
- 抽出した音源信号を送信して伝送量を削減

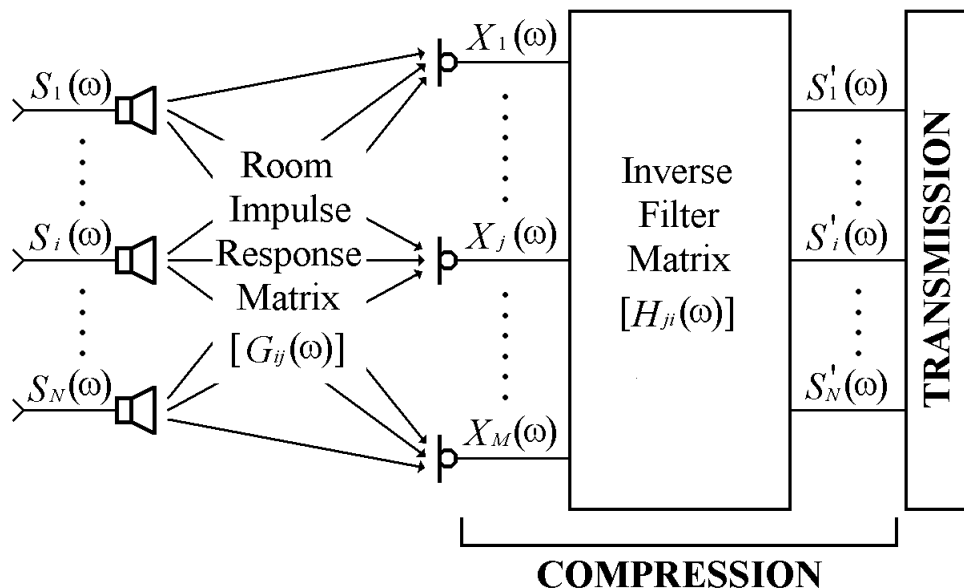
今までの研究成果

- 室内伝達関数の逆フィルタを用いて音源信号を抽出する空間符号化方式を提案
- 符号化による方向感への影響
 - 主観評価によって検討
 - 符号化による影響は見られず

今回の報告内容

- 符号化による「拡がり感」への影響を評価
 - 拡がり感
 - 音像が空間の広い領域に広がって知覚される現象
 - 両耳間相関係数によって客観的に評価
- 実用化に向けての更なるアプローチ
 - 前回までの符号化遅延...0.68sec
 - 符号化遅延がより短い場合の検討

符号化 (圧縮) : チャンネル信号に逆フィルタを畳み込み



$S_i(\omega)$: 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

$G_{ij}(\omega)$: 室内伝達関数

$$[S_i(\omega) \quad X_j(\omega)]$$

$H_{ji}(\omega)$: 逆フィルタ

$$[X_j(\omega) \quad S'_i(\omega)]$$

$\mathbf{G} = [G_{ij}(\omega)]$: 室内伝達関数行列 ($N \times M$)

$\mathbf{H} = [H_{ji}(\omega)]$: 逆フィルタ行列 ($M \times N$)

$$\mathbf{GH} = \mathbf{D}$$

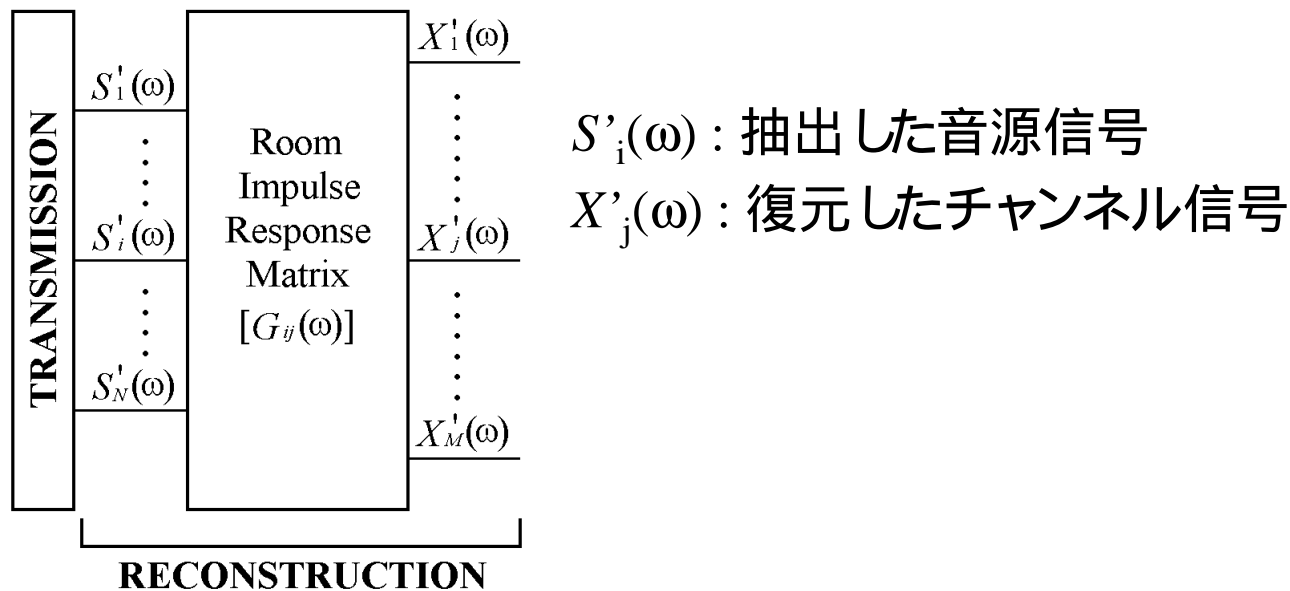
$$\mathbf{H} = \mathbf{G}^+ \mathbf{D}$$

\mathbf{D} : 単位行列 ($N \times N$)

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

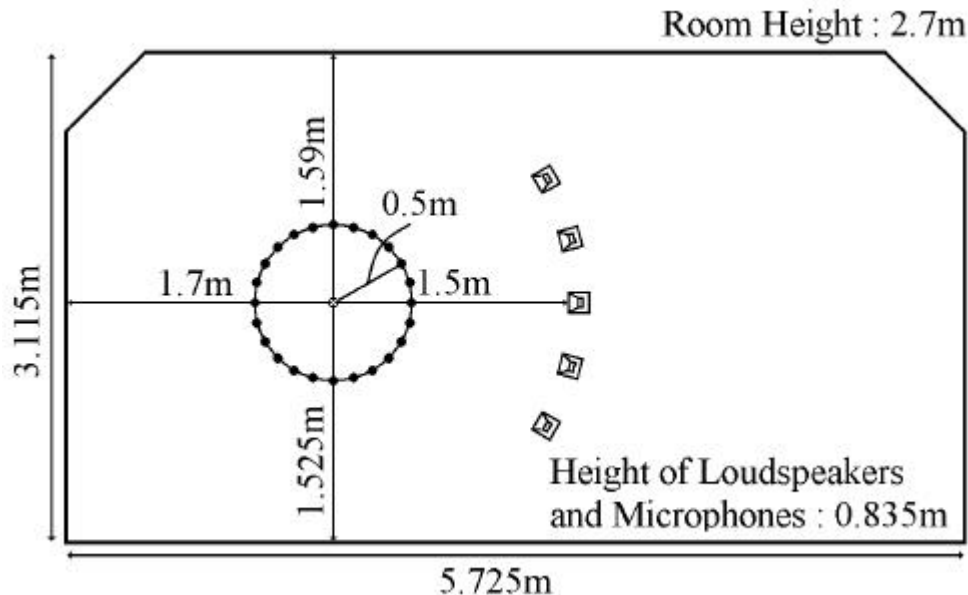
復号化 (復元)

抽出した音源信号に室内伝達関数を畳み込み
(送信側と受信側で室内伝達関数の情報は共有)



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)

室内伝達関数の測定 & チャンネル信号の作成



室内伝達関数の測定条件

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)
サンプリング周波数	48kHz	
参照信号	TSP(65536点)	
同期加算	8	
FIRフィルタ次数	7200	14400

*1 スピーカーから1m

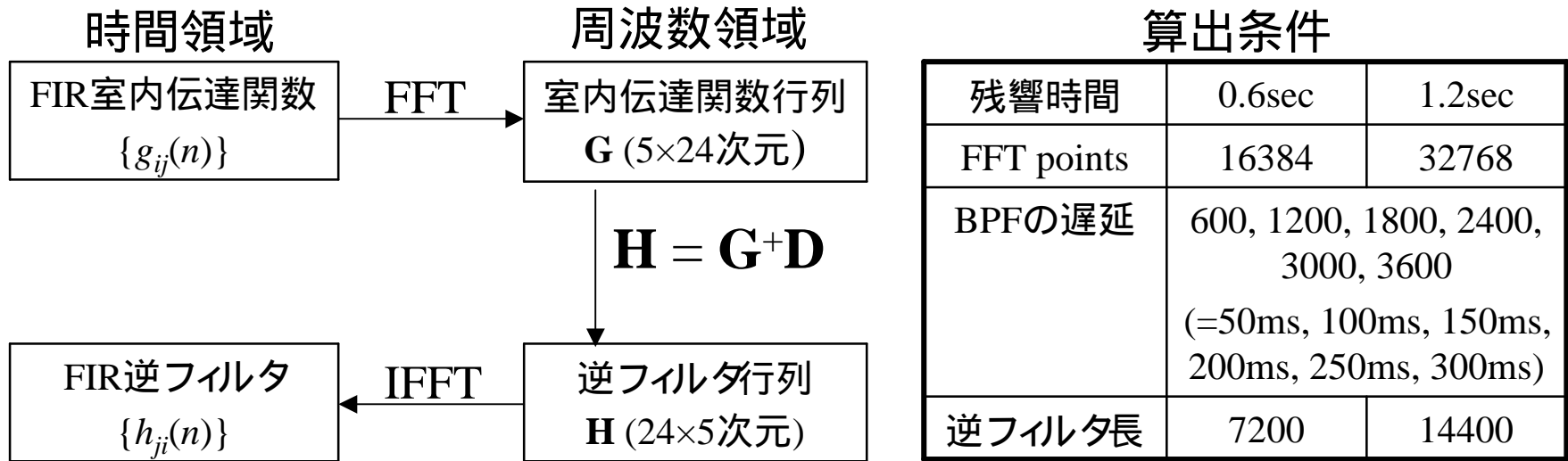
音源信号

種類	白色雑音	音声
個数	5	
長さ	5sec	
サンプリング周波数	12kHz	
帯域	50Hz ~ 5kHz	
備考	5個の音源は互いに無相関	

チャンネル信号の残響時間

150ms & 300ms $\xrightarrow{\times 4}$ 0.6sec & 1.2sec

逆フィルタの算出



D : FIRバンドパスフィルタのFFT変換

(サンプリング周波数12kHz、帯域50Hz ~ 5kHz)

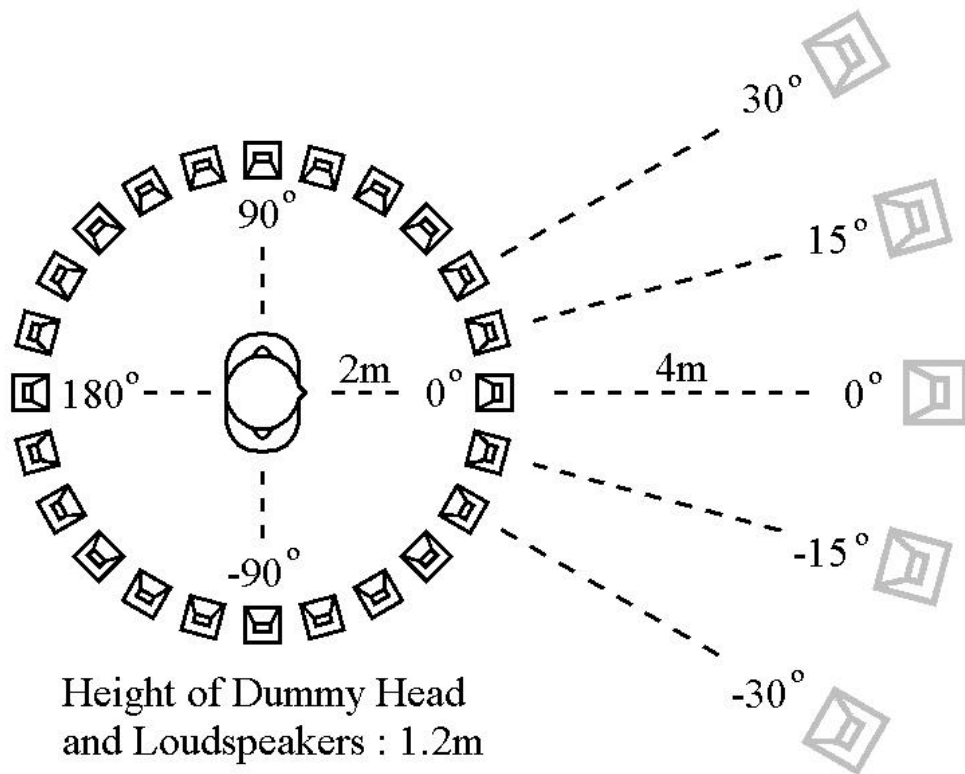
FIRバンドパスフィルタの遅延

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

両耳間相関係数(IACC)の算出には両耳の音圧信号が必要



頭部伝達関数の測定条件

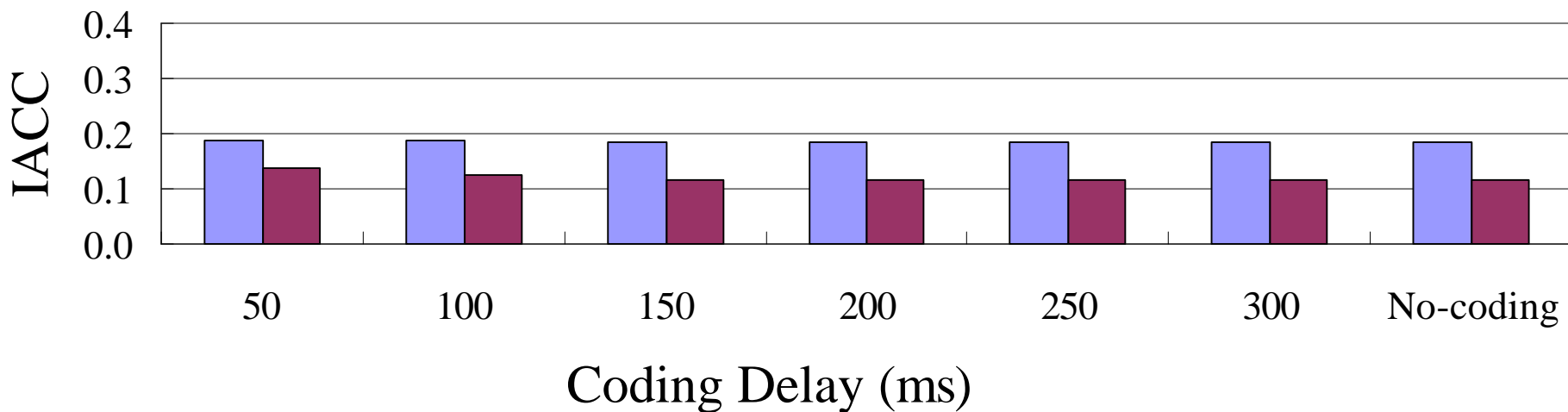
残響時間	80ms
暗騒音レベル	25.3dB(A)
音圧レベル*1	84.7dB(A)
サンプリング周波数	48kHz
参照信号	TSP(65536点)
同期加算	8
FIRフィルタ長*2	512

*1 ダミーヘッド位置

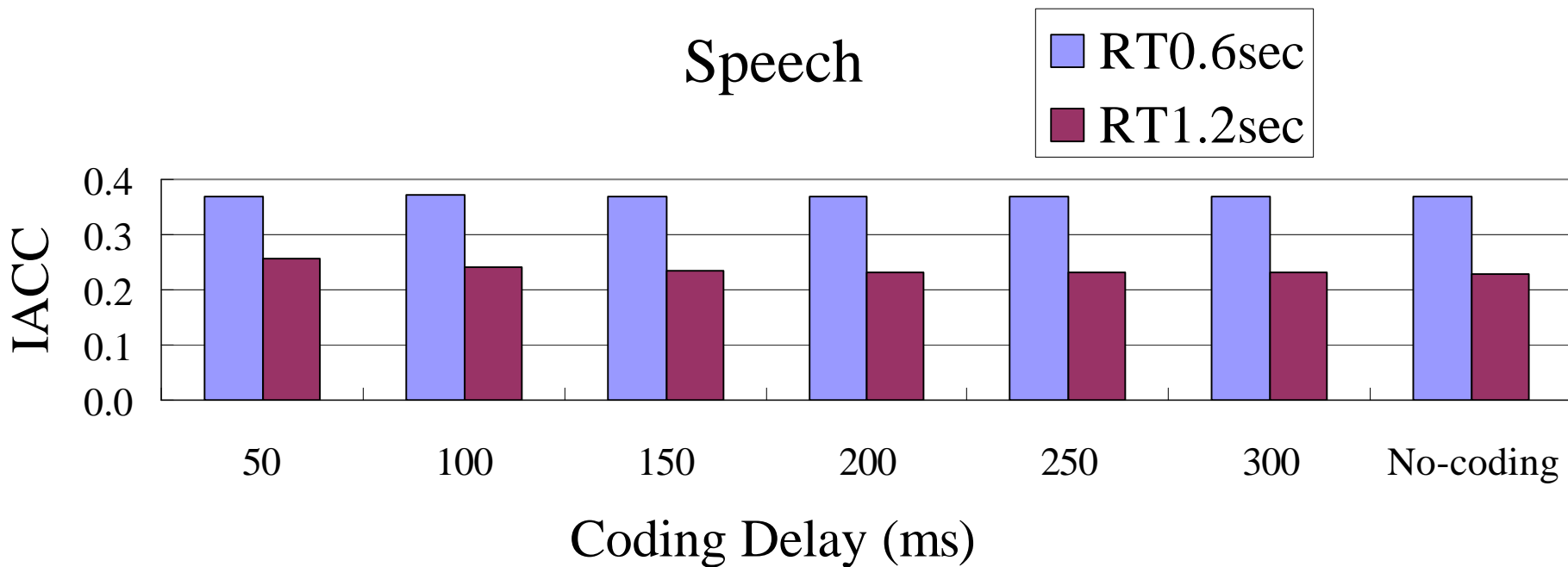
*2 直接波成分のみ切り出し

頭部伝達関数(HRTF)を測定し, 12kHzにダウンサンプリング
チャンネル信号にHRTFを畳み込み, 両耳の音圧信号を算出
両耳の音圧信号からIACCを算出(フレーム長...5sec)

White Noise



Speech



まとめ

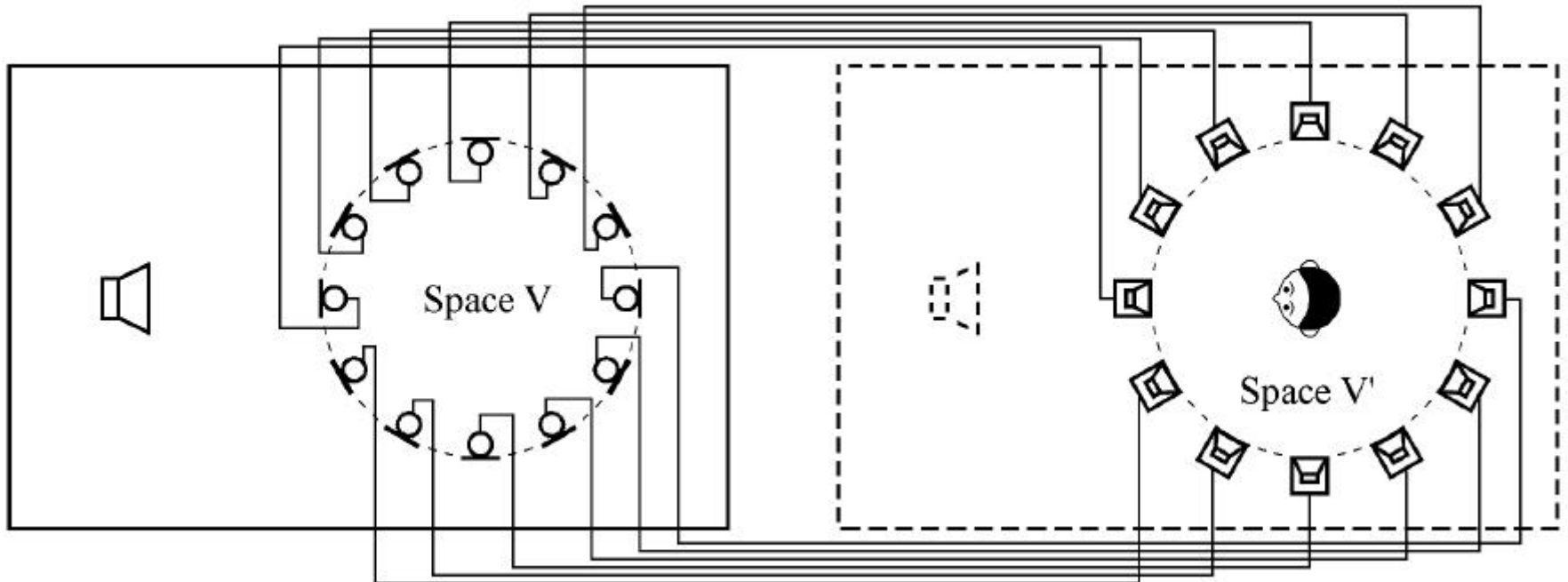
- 逆フィルタを用いた空間符号化方式による
拡がり感への影響を検討
- 両耳間相関係数によって拡がり感を検討
 - 符号化遅延が大きくなると拡がり感への影響
は見られなくなる
- 符号化遅延の小さい時
 - 拡がり感があまり変化していない
 - 実用化の際に非常に有効

今後の予定

- 主観評価による評価
 - 拡がり感, 空間印象
- 移動音源の場合の検討
- より広帯域な場合の検討

音場再現の原理

- ホイヘンスの原理
 - 空間 V の音場を空間 V' で再現
 - 再生部屋は無響室



両耳間相関係数の定義式

$$\text{IACC} = \frac{\lim_{T \rightarrow \infty} \frac{1}{2T} \sum_{n=-T}^T f_l(n) f_r(n - \tau_0)}{\sqrt{\lim_{T \rightarrow \infty} \frac{1}{2T} \sum_{n=-T}^T \{f_l(n)\}^2} \sqrt{\lim_{T \rightarrow \infty} \frac{1}{2T} \sum_{n=-T}^T \{f_r(n)\}^2}}$$

$$\tau_0 = \arg \max_{\tau} \left[\lim_{T \rightarrow \infty} \frac{1}{2T} \sum_{n=-T}^T f_l(n) f_r(n - \tau) \right]$$

τ_0 : 両耳間時間差(ITD)

$f_l(n)$: 左耳の音圧信号

$f_r(n)$: 右耳の音圧信号

