

3-P-29 位置情報による移動音源再現手法の主観的影響

◎木村敏幸(名大・情報科学), 箕一彦(中京大・情報科学), 武田一哉(名大・情報科学), 板倉文忠(名城大・理工)

1. はじめに

波面合成法による音場再現

- Huygensの原理により領域内の音場再現が可能
- 必要なチャンネル信号の数は非常に多い

遠隔システムの実現の際には
伝送量の削減が必要

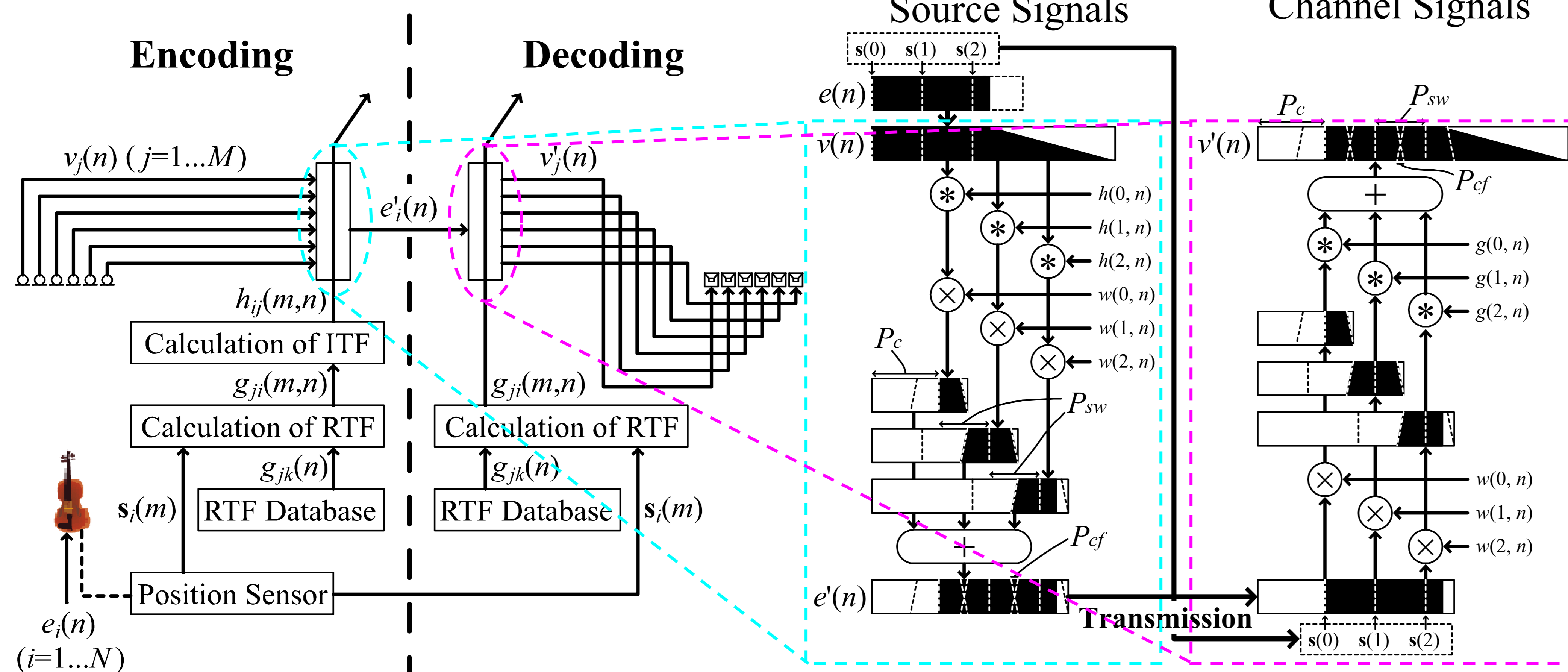
音源抽出に基づく空間符号化手法

- 伝送量: チャンネル信号の数 → 音源の数
- 今までの検討: 静止音源を対象

移動音源の抽出に基づいた
空間符号化手法の提案

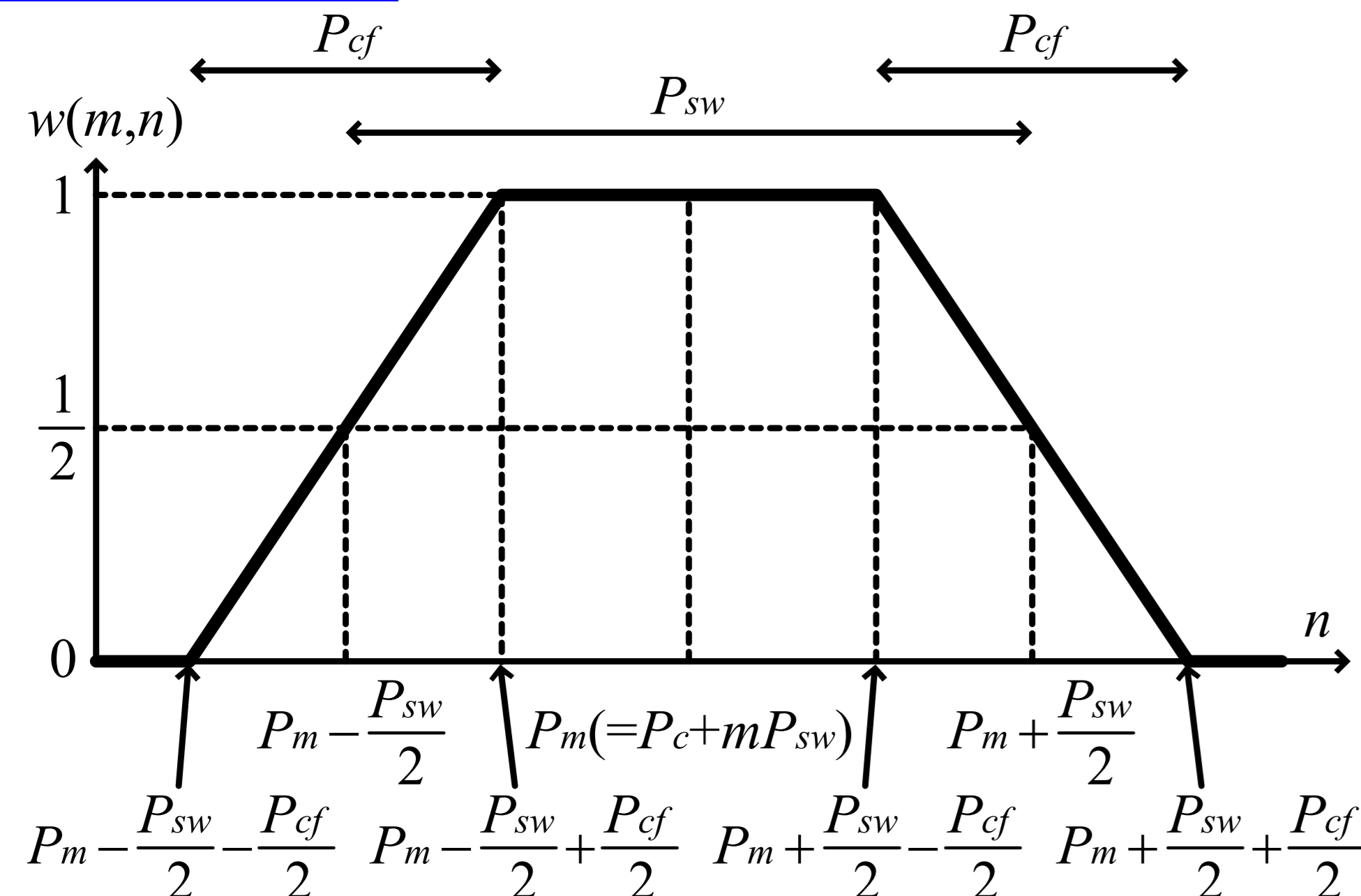
2. アルゴリズム

提案手法のブロック図



- $e_i(n)$: i 番目の音源信号
- $v_j(n)$: j 番目のチャンネル信号
- $s_i(m)$: i 番目の音源の位置情報
- $g_{ji}(m, n)$: 時間 m における室内伝達関数
- $h_{ij}(m, n)$: 時間 m における逆伝達関数
- m, n : 離散時間(サンプリング周波数 F_p, F_s)

窓関数の形状



音源信号の抽出

$$e'_i(n) = \sum_{m=0}^{\infty} w(m, n) \sum_{j=1}^M h_{ij}(m, n) * v_j(n)$$

- $e'_i(n)$: 抽出した i 番目の音源信号

逆伝達関数の算出

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

$$\mathbf{H}(m, \omega) = [H_{ij}(m, \omega)] = [FFT\{h_{ij}(m, n)\}]$$

$$\mathbf{G}(m, \omega) = [G_{ji}(m, \omega)] = [FFT\{g_{ji}(m, n)\}]$$

$$\mathbf{D}(\omega) = \text{diag}(e^{-j\omega P_c}, \dots, e^{-j\omega P_c})$$

- $\mathbf{G}^+(m, \omega)$: $\mathbf{G}(m, \omega)$ のMoore-Penrose擬似逆行列
- $P_c (=F_s T_c)$: 因果律を満たすための遅延サンプル数

チャンネル信号の復元

$$v'_j(n) = \sum_{m=0}^{\infty} \sum_{i=1}^N g_{ji}(m, n) * [w(m, n) e'_i(n)]$$

- $v'_j(n)$: 復元した j 番目のチャンネル信号

3. 符号化実験

チャンネル信号の作成

- 虚像法を用いて計算機上で作成

$$v_j(n) = \sum_{i=1}^N \sum_{p_x=0}^1 \sum_{p_y=0}^1 \sum_{q_x=-\infty}^{\infty} \sum_{q_y=-\infty}^{\infty} D[\mathbf{d}_{ji}^{pq}(n+a)] \beta^{q_x - p_x + |p_x| + q_y - p_y + |p_y|} \frac{e_i[n+a - \frac{F_s}{c} |\mathbf{d}_{ji}^{pq}(n+a)|]}{4\pi |\mathbf{d}_{ji}^{pq}(n+a)|}$$

$$\mathbf{d}_{ji}^{pq}(n+a) = \mathbf{s}_{ji}^{pq}(n+a) - \mathbf{r}_j = \begin{bmatrix} (1-2p_x)s_{jx}(n+a) + 2q_x L_x \\ (1-2p_y)s_{jy}(n+a) + 2q_y L_y \end{bmatrix} - \begin{bmatrix} r_{jx} \\ r_{jy} \end{bmatrix}$$

- c : 音速(=340[m/s])
- $\mathbf{s}_{ji}^{pq}(n+a)$: i 番目の音源の虚像位置ベクトル
- \mathbf{r}_j : j 番目の受信点位置ベクトル
- $D[\mathbf{d}_{ji}^{pq}(n+a)]$: 受信点の指向性関数(カーゴイド型)
- \mathbf{u}_j : j 番目の受信点の指向ベクトル

チャンネル信号の作成条件

ドライソース	音声	フルート
F_s (サンプリング周波数)	48[kHz]	
音源の長さ	4[second]	
β (壁面の反射係数)	0.5	0.7
最大の壁面反射回数	6	10
残響時間[second]	0.6	1.0
V (音源の移動速度)	1[m/s]=3.6[km/h]	

室内伝達関数の算出

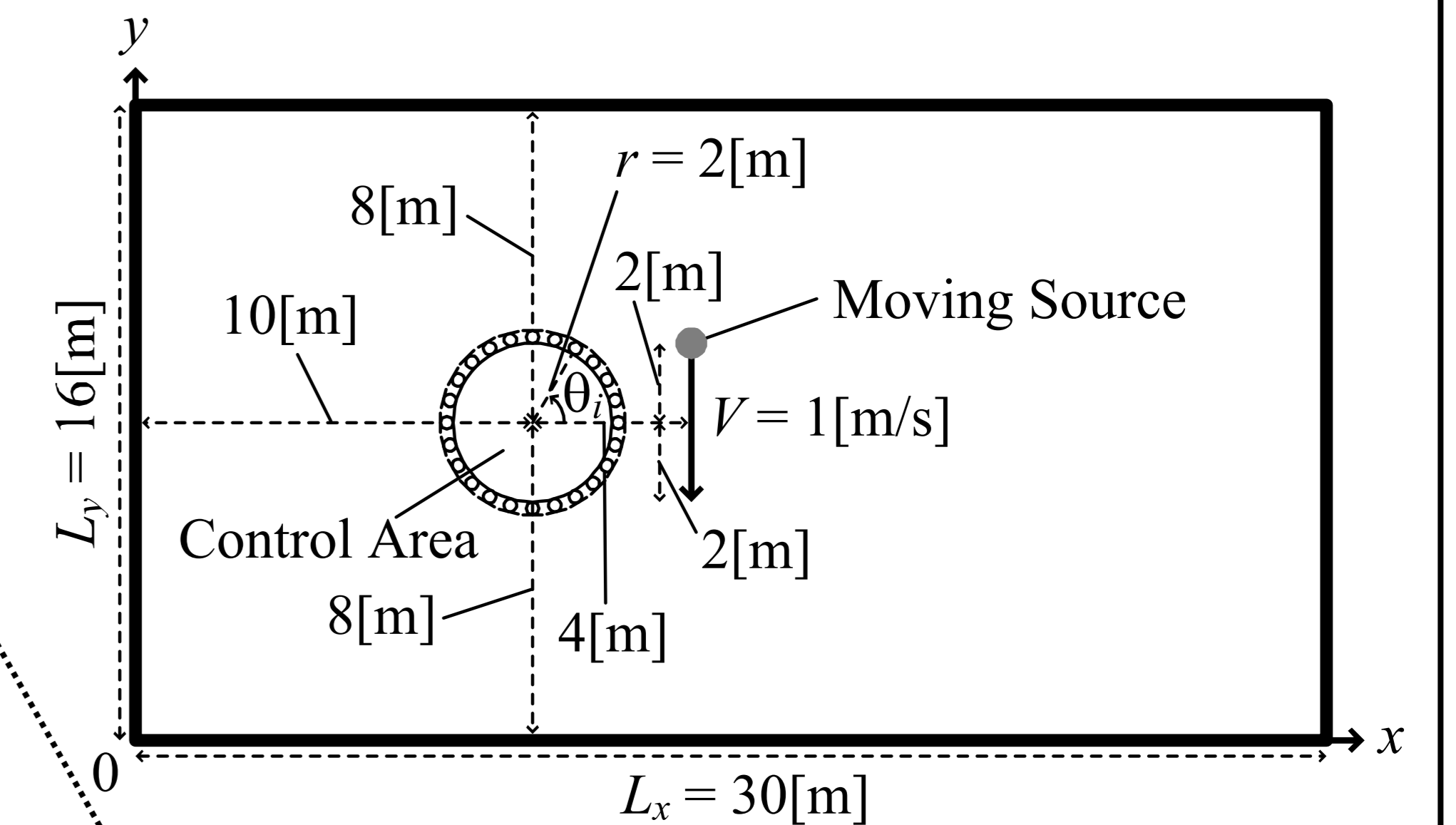
- $\mathbf{s}_k(k=1 \dots 481)$ から $\mathbf{r}_j(j=1 \dots 24)$ までの室内伝達関数を算出

$$g_{jk}(n) = \sum_{p_x=0}^1 \sum_{p_y=0}^1 \sum_{q_x=-\infty}^{\infty} \sum_{q_y=-\infty}^{\infty} D[\mathbf{d}_{jk}^{pq}] \beta^{q_x - p_x + |p_x| + q_y - p_y + |p_y|} \frac{\delta[n - \text{round}(\frac{F_s}{c} |\mathbf{d}_{jk}^{pq}|)]}{4\pi |\mathbf{d}_{jk}^{pq}|}$$

$$\mathbf{u}_j = \begin{pmatrix} \cos \frac{\pi(j-12)}{12} \\ \sin \frac{\pi(j-12)}{12} \end{pmatrix}, \mathbf{s}_k = \begin{pmatrix} 14 \\ 10 - \frac{k-1}{120} \end{pmatrix}, \mathbf{r}_j = 2\mathbf{u}_j + \begin{pmatrix} 10 \\ 8 \end{pmatrix}$$

- 時間 m における室内伝達関数を算出

$$g_{ji}(m, n) = g_{jk}(n) \quad (m = \frac{F_p(k-1)}{120})$$



- 時間 $n+a$ に信号を入力すると、時間 $n+a + \frac{F_s}{c} |\mathbf{d}_{ji}^{pq}(n+a)|$ にチャンネル信号が出力
- $a(0 \leq a < 1)$ は n ごとに以下の二次方程式により算出

$$\left\{ \Delta \mathbf{d}_{ji}^{pq}(n) \right\}^2 - \frac{c^2}{F_s^2} a^2 + 2 \left\{ \mathbf{d}_{ji}^{pq}(n) \cdot \Delta \mathbf{d}_{ji}^{pq}(n) \right\} + \frac{c^2}{F_s^2} (K-n)a + \left\{ \mathbf{d}_{ji}^{pq}(n) \right\}^2 - \frac{c^2}{F_s^2} (K-n)^2 = 0$$
- $\Delta \mathbf{d}_{ji}^{pq}(n) = \mathbf{d}_{ji}^{pq}(n+1) - \mathbf{d}_{ji}^{pq}(n)$
- K : 以下の条件を満たす任意の整数

$$\text{ceil}(n + \frac{F_s}{c} |\mathbf{d}_{ji}^{pq}(n)|) \leq K \leq \text{floor}(n+1 + \frac{F_s}{c} |\mathbf{d}_{ji}^{pq}(n+1)|)$$
- 入力信号 $e_i(n+a)$ は補間を行ってから適用

$$e_i(n+a) = (1-a)e_i(n) + ae_i(n+1)$$

逆伝達関数の算出

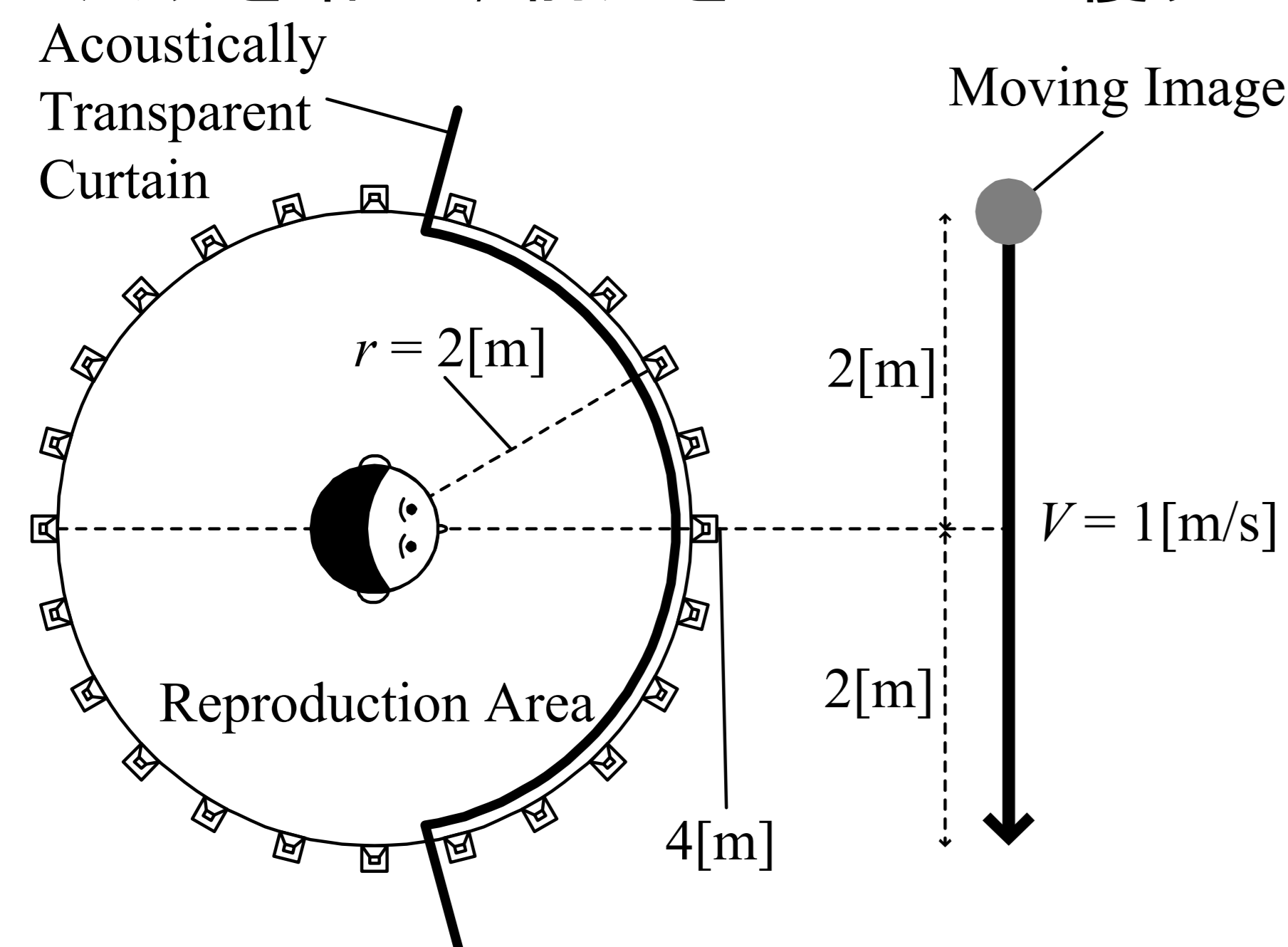
逆伝達関数の算出条件

残響時間[second]	0.6	1.0
FFTフレーム長[sample]	65536	131072
符号化遅延時間 T_c ($=P_c = T_c F_s$)	20[ms](=960[sample])	
フィルタ長[sample]	28800	48000
サンプリング周波数 F_p ($=P_{sw} = F_s / F_p$)	30, 60, 120[Hz] (=1600, 800, 400[sample])	
クロスフェード時間 T_{cf} ($=P_{cf} = T_{cf} F_s$)	1, 4[ms] (=48, 192[sample])	

4. 主観評価実験

実験環境

- 残響時間...約80ms
- 暗騒音レベル...25.0dB(A)
- 音圧レベル...受聴者位置で約70dB(A)
- 照明を暗くし、前方をカーテンで覆う



実験手順

- 前半(動き評価)
 - X...基準となる動き方をする音
 - A or B...どちらかは必ず基準の音と同じ動き方
 - 同じ動き方だと思った方を「5.0」と評定
 - もう一方は表に従って「1.0」から「4.9」までで評定
- 前半(全体評価)
 - X...原音
 - A or B...どちらかは必ず原音と同じ音
 - 原音と思った方を「5.0」と評定
 - もう一方は表に従って「1.0」から「4.9」までで評定

評定表

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

実験計画

- 被験者...男子学生8名
- 実験手法...XAB法
- 練習試行...12
 - 6(符号化音の種類)×2(A or Bに割り当て)
- 本試行...24
 - 6(符号化音の種類)×2(A or Bに割り当て)×2(繰り返し)

Subjective Assessment

Moving Evaluation	Total Evaluation
Session 1 Order...Randomized (Speech or Flute)	Session 2 Order...Randomized (Speech or Flute)

Evaluation

Session	Trial
Practice (12 trials)	Main (24 trials)

X	Break	A	Break	B
Ref.	(0.5sec)	Ref./Test	(0.5sec)	Test/Ref.

符号化音の種類

	1	2	3	4	5	6
F_p	30Hz	30Hz	60Hz	60Hz	120Hz	120Hz
T_{cf}	1ms	4ms	1ms	4ms	1ms	4ms

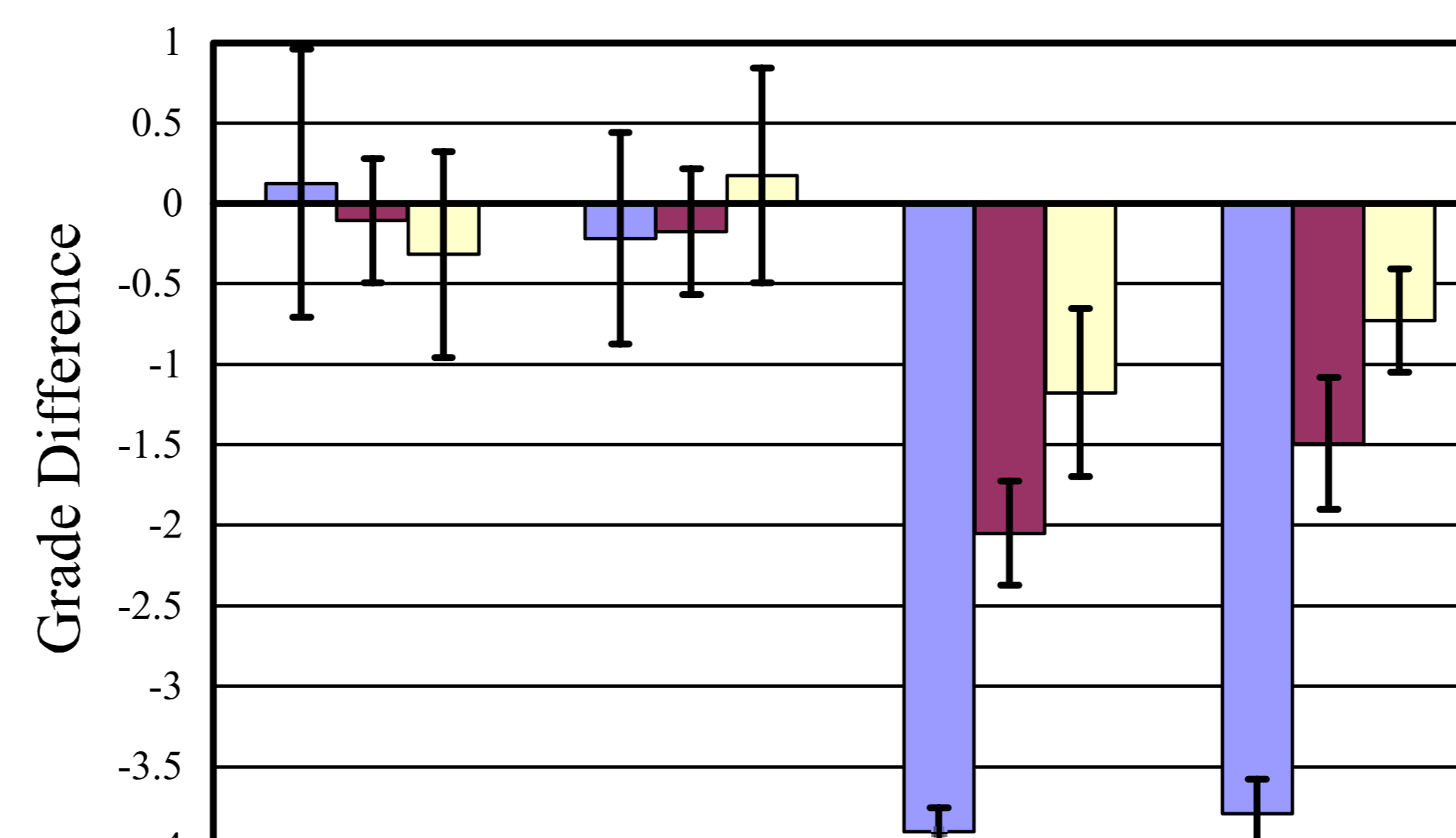
被験者の選定

- 原音を割り当てた刺激を「5.0」と評定した回数を検査
- 各セッションにおける成績上位3名の結果を分析に使用

実験結果

- フルート
 - 音声よりも全体的に低い評定差値
 - $F_p=120$ [Hz]で十分実用的なシステムを構築可能
- 音声
 - $F_p=30$ [Hz]でも実用的なシステムを構築可能

Speech, $T_{cf}=1$ [ms] Speech, $T_{cf}=4$ [ms] Flute, $T_{cf}=1$ [ms] Flute, $T_{cf}=4$ [ms]

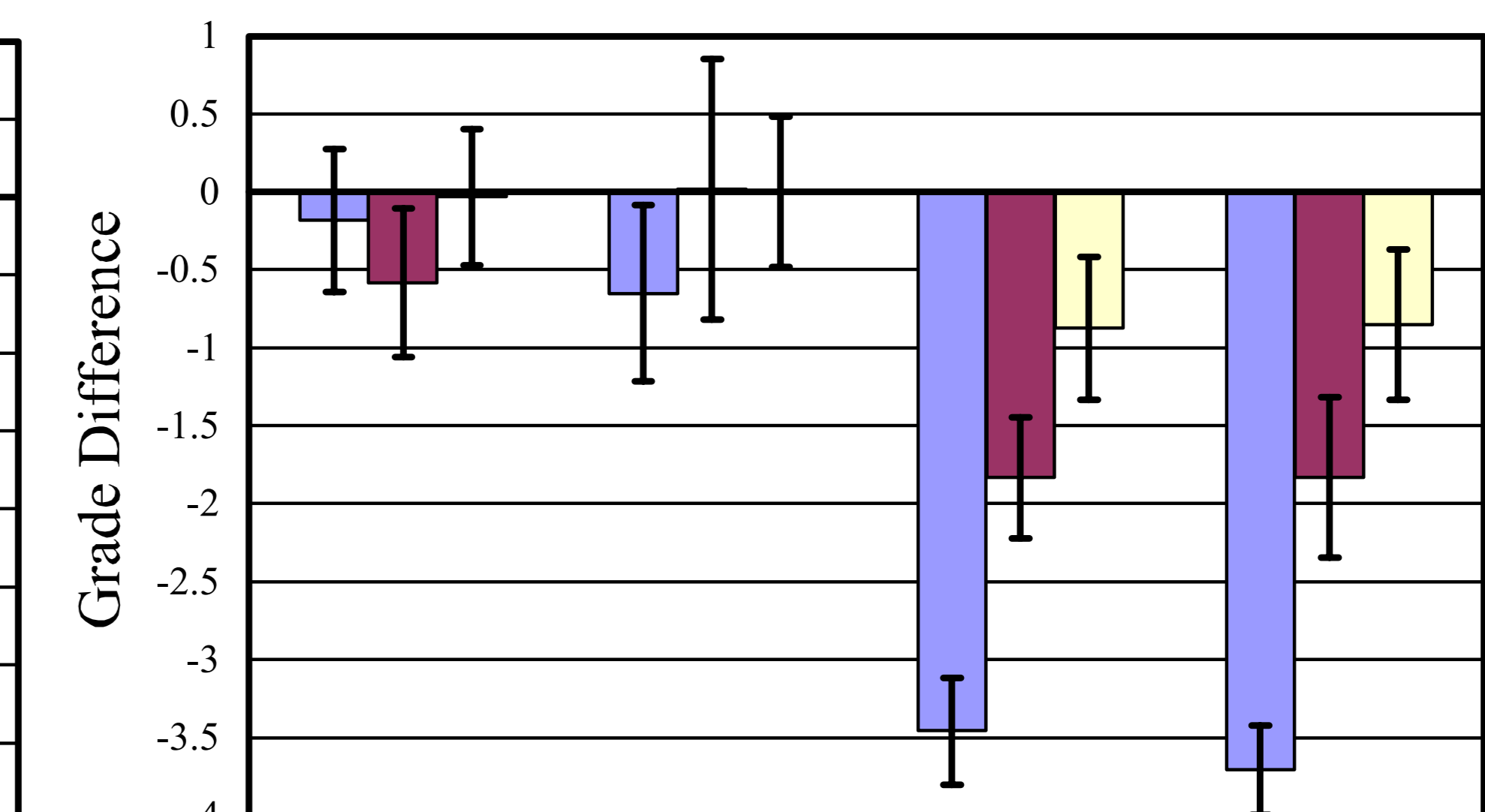


動き評価の結果

各被験者の識別結果

被験者	標本数	動き評価		全体評価	
		音声	フルート	音声	フルート
A	24	15	23	15	22
B	24	7	23	16	24
C	24	7	24	14	24
D	24	14	21	11	21
E	24	14	14	15	17
F	24	11	22	9	22
G	24	11	19	9	20
H	24	9	18	10	17

Speech, $T_{cf}=1$ [ms] Speech, $T_{cf}=4$ [ms] Flute, $T_{cf}=1$ [ms] Flute, $T_{cf}=4$ [ms]



全体評価の結果

5. まとめ

- 音源の位置情報を基に移動音源を抽出することで伝送量を削減する空間符号化手法を提案
 - 虚像法によって残響空間上のチャンネル信号を模擬し、符号化実験を実行
 - 実験によって達成された伝送量の削減:チャンネル信号の数(24チャンネル)→音源の数(1チャンネル)
- 符号化性能を評価するために、主観評価実験を実施
 - システムの対象となる音源に対して適切なパラメータを設定すれば、知覚への影響がほとんどないシステムを構築可能