



National Institute of Information and Communications Technology

# Multiple Vertical Panningを用いた 立体音響システムのための 聴取実験

2012/3/16

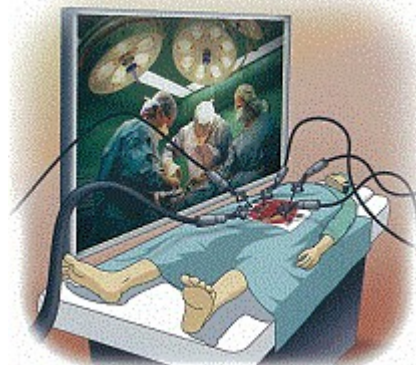
木村敏幸, 安藤広志

独立行政法人 情報通信研究機構(NICT)

ユニバーサルコミュニケーション研究所

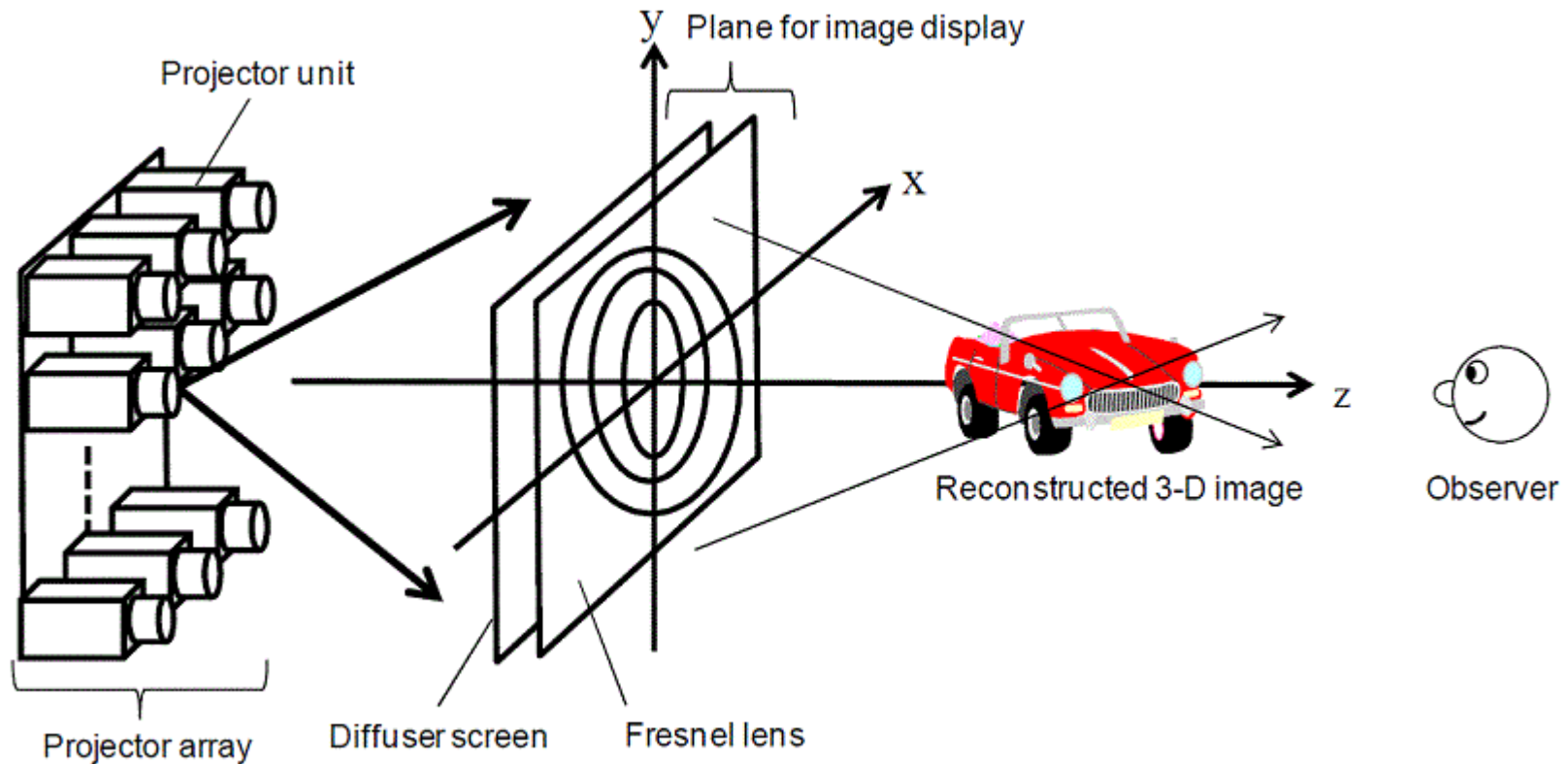
# 超臨場感コミュニケーション

- 立体映像技術や立体音響技術によって映像や音響をよりリアルに表現
- より臨場感のあるコミュニケーションが可能
  - 立体テレビや立体遠隔通信会議など



# 大画面立体映像提示システム

- プロジェクタアレイを使用
  - 拡散スクリーンで垂直方向のみに光を拡散
  - 水平方向にのみ視差付き映像を提示
- 水平位置に応じた視差映像が鑑賞可能



# 大画面立体映像用音響システムの技術要件

- 複数の人がどこで聞いても対象物の位置から音が出ているように聞こえること
  - ステレオホニック, 5.1chシステムは適用が困難
- 何も装着せずに音が聞けること
  - ヘッドホンを使ったシステムは適用が困難
- プロジェクタアレイとスクリーンの間には何も置かないこと
  - 22.2chシステム, 高次アンビソニック, 波面合成法は適用が困難
- 会議システムとして応用する場合, マイクロホンも間におかないこと
  - トランスオーラル, 境界音場制御は適用が困難

# 本研究の目的

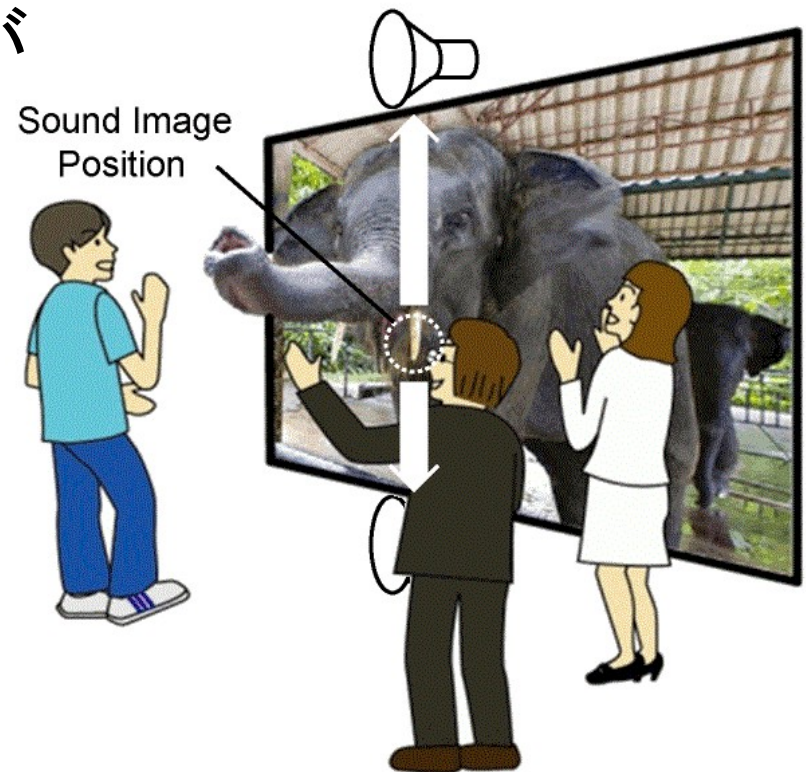
---

- 大画面立体映像提示システムに適合した立体音響システム
    - 今までの立体音響システムでは対応困難
- ↓
- 新たなる立体音響システムを提案
    - 今までの立体音響システムとは異なる観点
    - Multiple Vertical Panning (MVP)方式



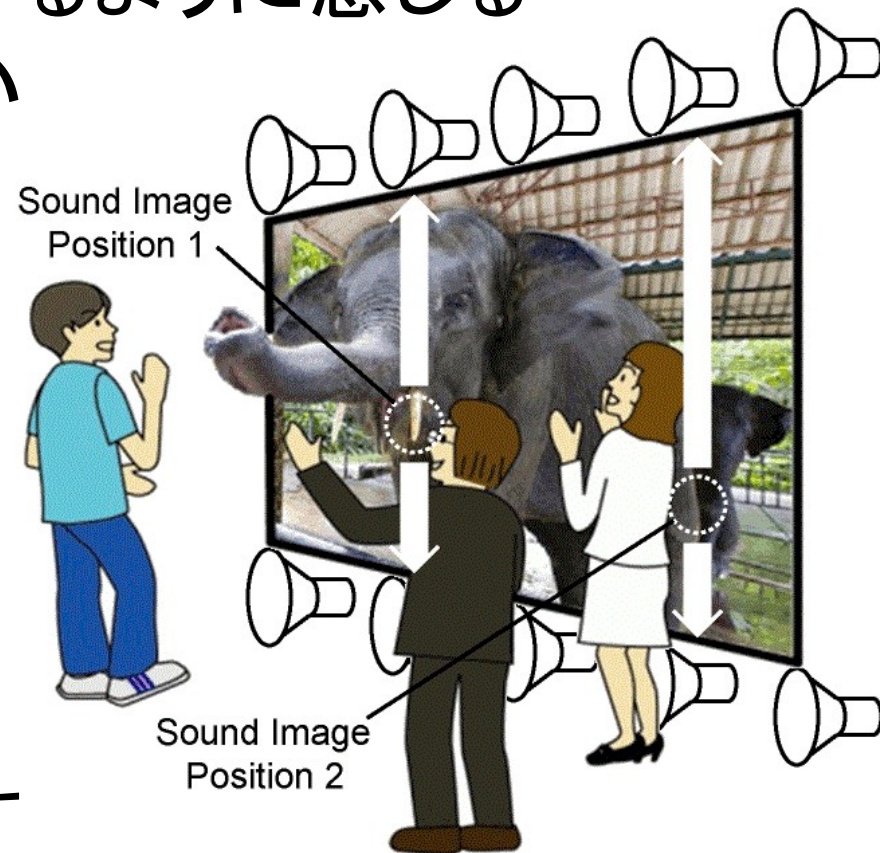
# 提案システムの原理

- 音像表現位置の上下にスピーカを2個配置
  - 立体像の描写位置(スクリーン)の上下に配置
- 音源に音量差をつけ2個のスピーカから音を再生(垂直パニング)
  - スピーカ2個の間で音が鳴っているように感じる
- 音量差を適切に設定
  - 複数の聴取者はどこにいても常に立体像の位置で音が鳴っているように感じる



# 提案システムの原理

- スクリーンの上下にスピーカ対を複数設置
  - 音像の表現位置が左右方向にも拡大
- 複数の聴取者が常にスクリーン上の立体像の位置で音が鳴っているように感じる
- ヘッドホンを利用しない
- プロジェクタアレイとスクリーンの間に何も置いていない
  - マイクロホンも自由に配置が可能



↓  
NICT 技術要件を全て満たす

- 提案システム
  - 最小構成要素
    - スクリーン上下に配置した2個のスピーカ
  - 聴覚性能
    - 最小の構成要素の重ね合わせで表現可能
  - スクリーン上下に配置した2個のスピーカによる合成音像の高さ知覚を評価すれば十分



- 2個のスピーカによる合成音像の高さ知覚を評価する聴取実験
  - 大画面裸眼立体映像提示システムのスクリーンの上下に配置することを想定
  - 垂直パニング曲線も合わせて検討



# 実験環境

- 可変残響室(ATRビル地下1階)
  - 残響時間: 140 ms, 1030 ms
  - 暗騒音: 14 dBA(140 ms), 22 dBA(1030 ms)

140ms



1030ms



# 実験環境

- スピーカを縦に27個配置

- スピーカアレイの高さ

- 2.97 m  
(=11 cm×27)

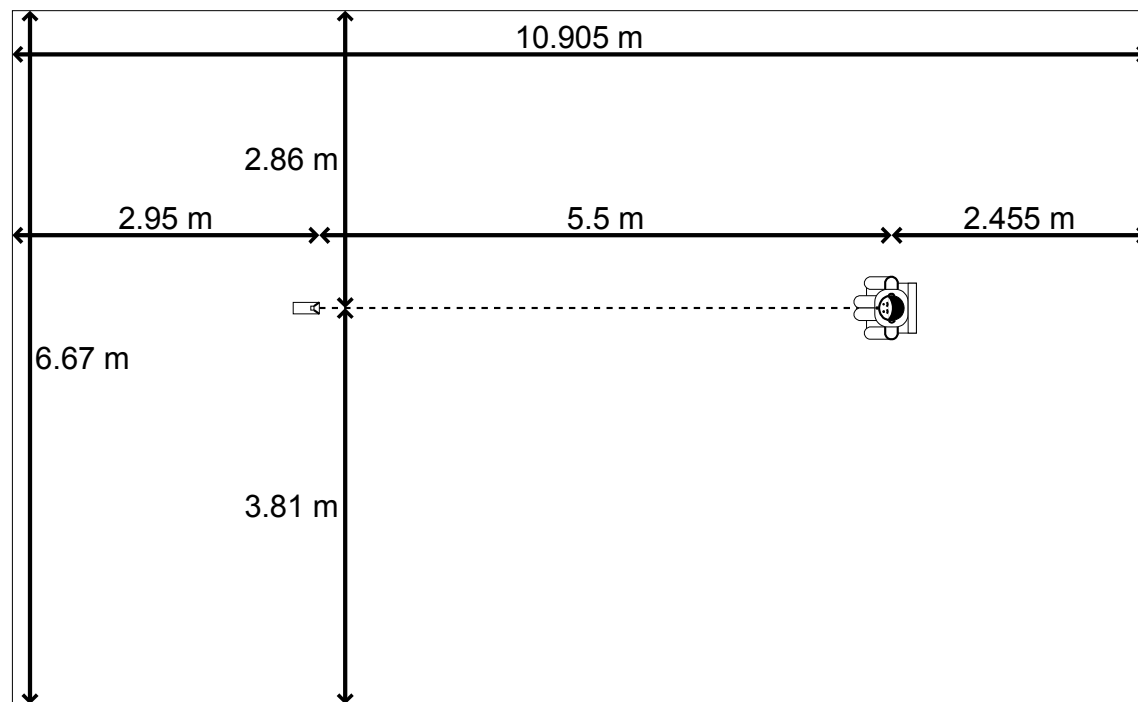
- 聴取位置

- スピーカアレイから5.5 m

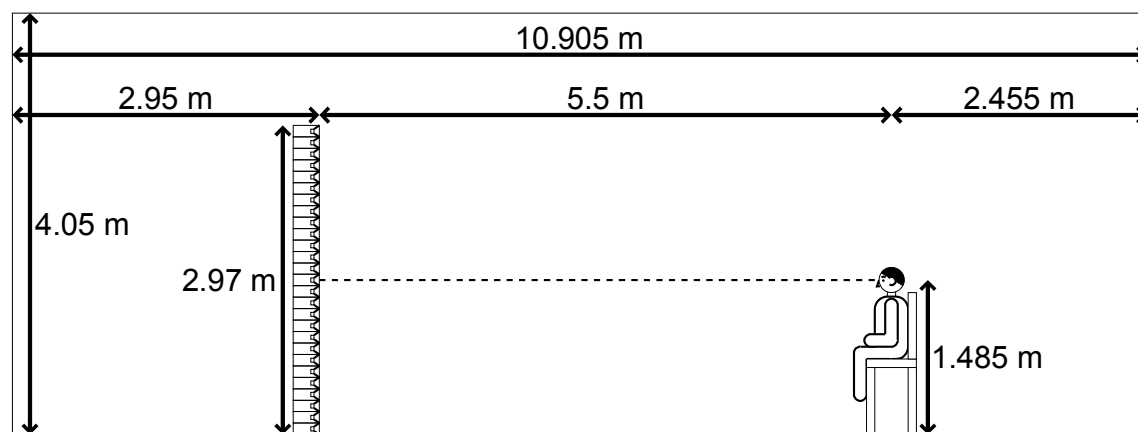
- 音圧レベル

- 聴取位置で約70 dBA

Plane View



Cross-sectional View



- パニング条件

- 最上部と最下部の計2個のスピーカから音量差のみをつけた音を再生
- 条件の数: 31条件(-15dB~15dBで1dB間隔)

(a) Panning Condition



- 統制条件

- 27個のスピーカのうち1個から音を再生
- 条件の数: 13条件

(b) Control Condition

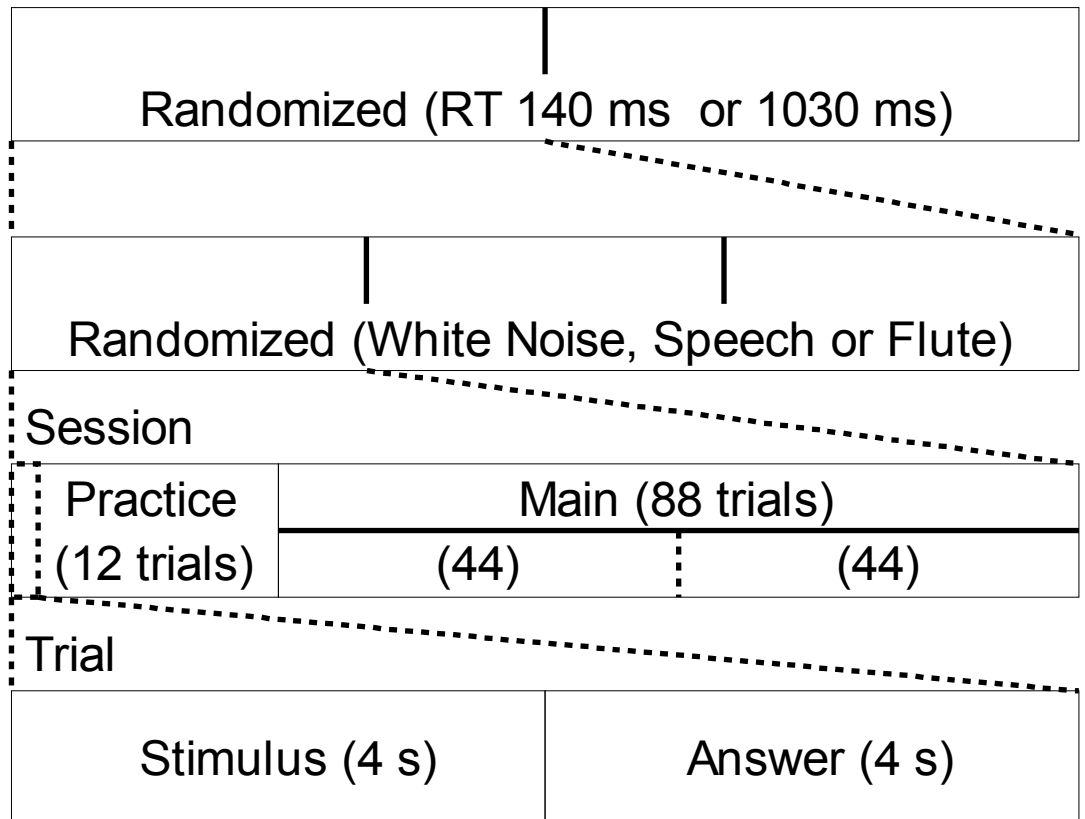


# 実験計画

- 音源
  - 白色雑音, 音声, フルート
  - 4秒間

- 被験者
  - 12名
    - 男性6名
    - 女性6名
  - 年齢
    - 21~32歳
  - 聴力
    - 会話域正常

## Listening Test

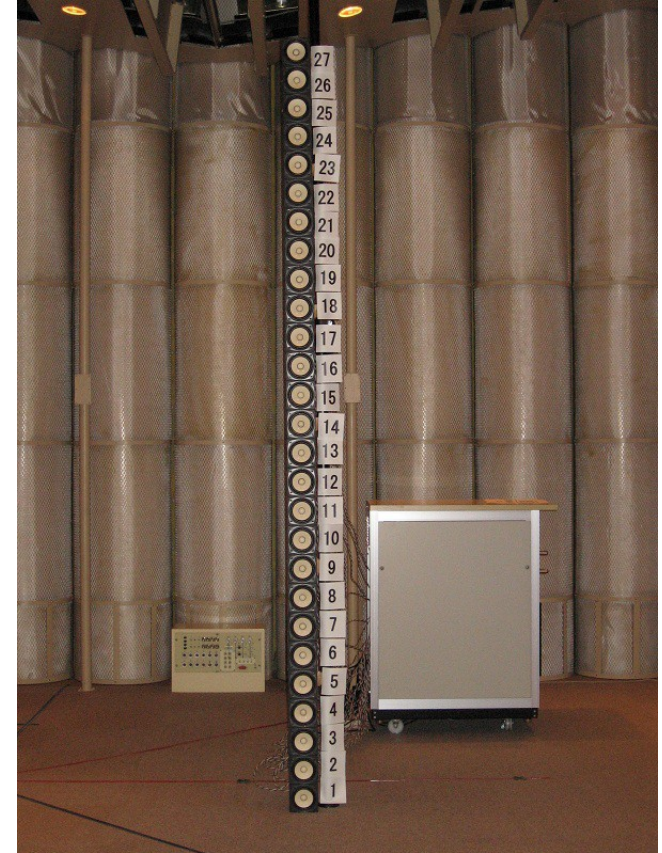




- セッション数
  - $6 = \text{音源}(3) \times \text{残響時間}(2)$
  - 音源, 残響時間の順番はランダムイズ
- 練習試行
  - $12 = (7(\text{パニング}) + 5(\text{統制条件})) \times 1(\text{繰り返し})$
  - 順番はランダムイズ
- 本試行
  - $88 = (31(\text{パニング}) + 13(\text{統制条件})) \times 2(\text{繰り返し})$ 
    - 44試行ごとに休憩を導入
  - 順番はランダムイズ

# 実験手順

- 27個のスピーカの中から音が再生されたと感じた番号を回答用紙に記入
  - スピーカ右側に指標を表示
  - 2つ以上音が鳴っていると思った場合は, 2つ以上の番号で回答
  - 頭部及び上半身の動きは許可



# 実験結果の分析

- スピーカ番号が複数記入された回答を除去
- 知覚した音像の高さを算出

$$H_{\text{per}}[\text{m}] = (I_{\text{ans}} - 14) \times 0.11$$

–  $I_{\text{ans}}$  : 回答したスピーカ番号

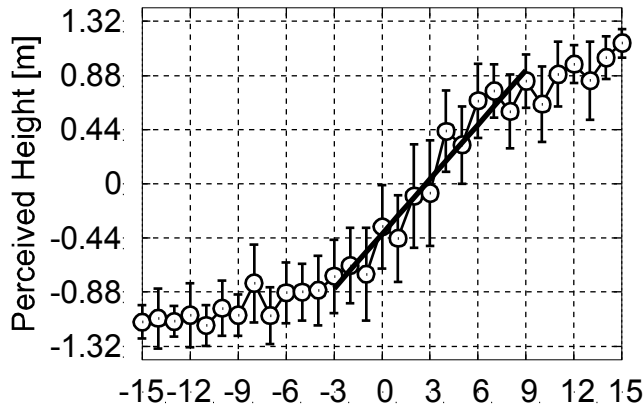
–  $H_{\text{per}}$  : 知覚した音像の高さ

- 1番 : -1.43 m
- 14番 : 0 m
- 27番 : 1.43 m

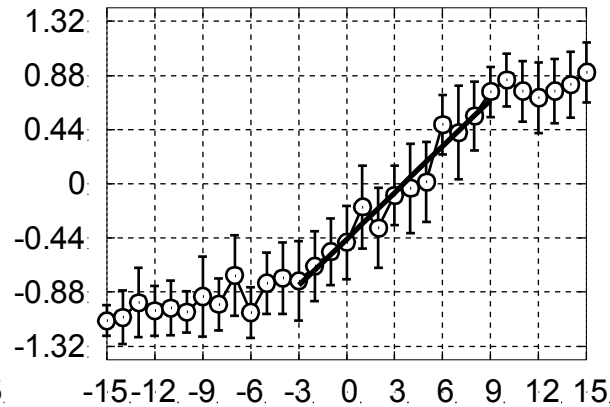
# 実験結果(パニング条件)

- 音量差が-3~9dBのときに直線的に移動

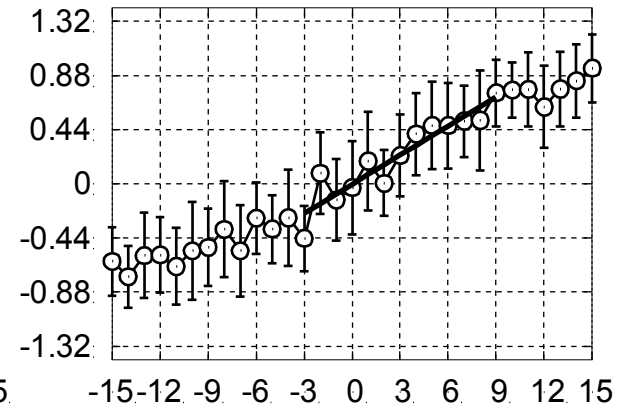
White Noise (140ms)



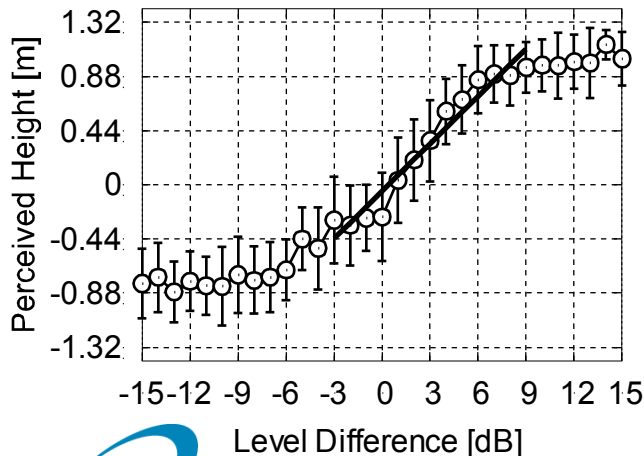
Speech (140ms)



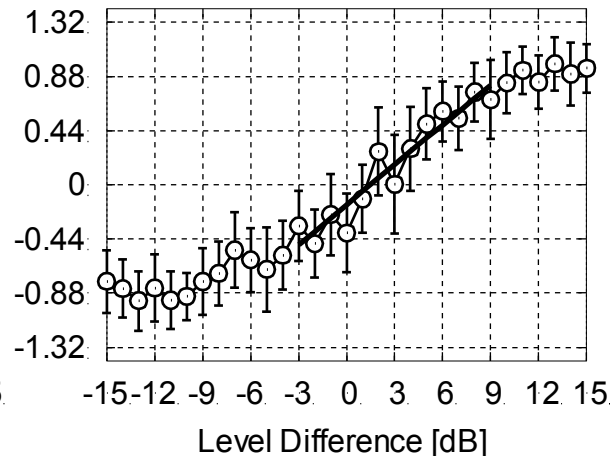
Flute (140ms)



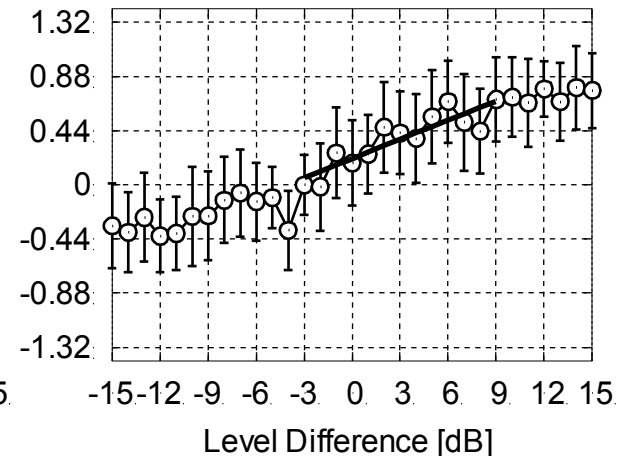
White Noise (1030ms)



Speech (1030ms)



Flute (1030ms)



# パニング曲線の算出

- 回帰直線の平均より算出

$$H_{\text{pan}} = \begin{cases} -1.32 & (\Delta A < -11.05) \\ 0.1065 \Delta A - 0.1437 & (-11.05 \leq \Delta A \leq 13.74) \\ 1.32 & (\Delta A > 13.74) \end{cases}$$

–  $\Delta A$ : 音量差

- 知覚する音の高さの弁別閾を算出

$$DL_{\text{pan}}^+ = \tan(\tan^{-1}(H_{\text{pan}}/5.5) + \phi) \times 5.5$$

$$DL_{\text{pan}}^- = \tan(\tan^{-1}(H_{\text{pan}}/5.5) - \phi) \times 5.5$$

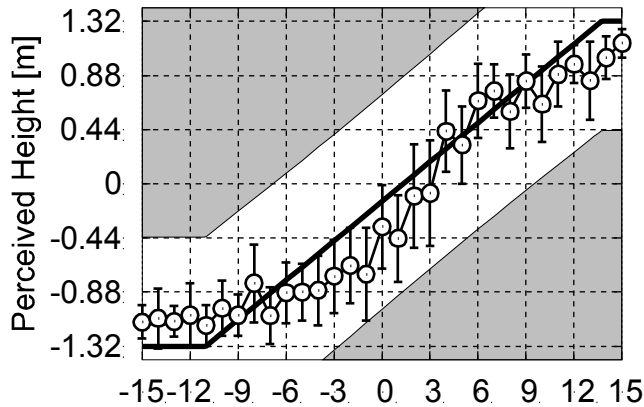
–  $\phi$  (=9度): 音の高さの弁別角



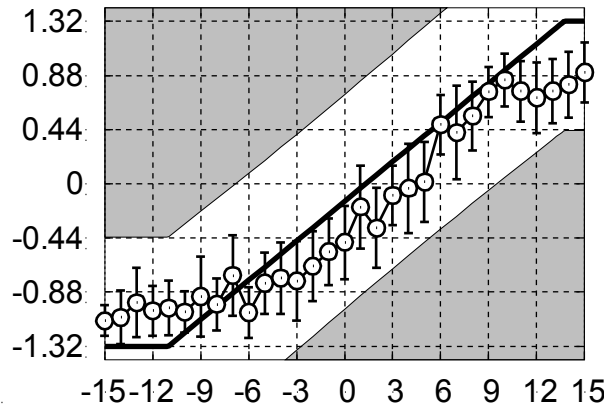
# 実験結果(パニング曲線)

- フルート, 残響時間1030ms以外弁別不能

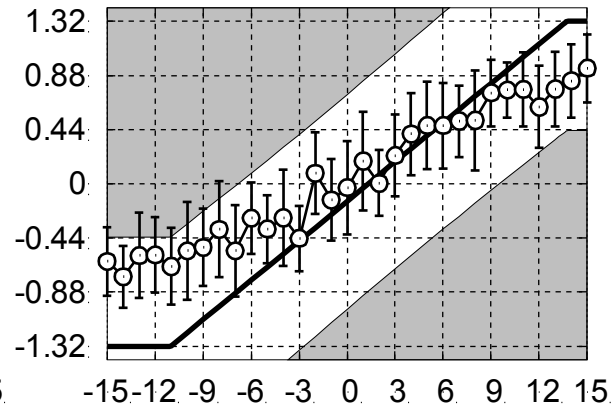
White Noise (140ms)



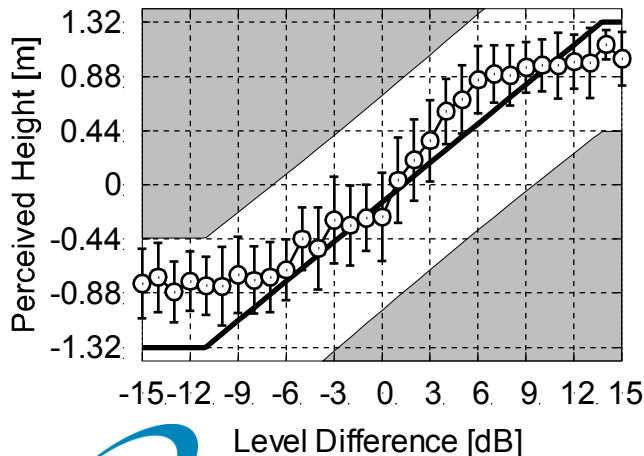
Speech (140ms)



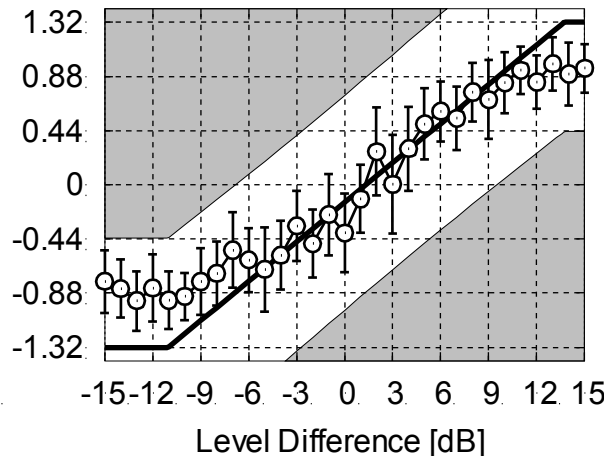
Flute (140ms)



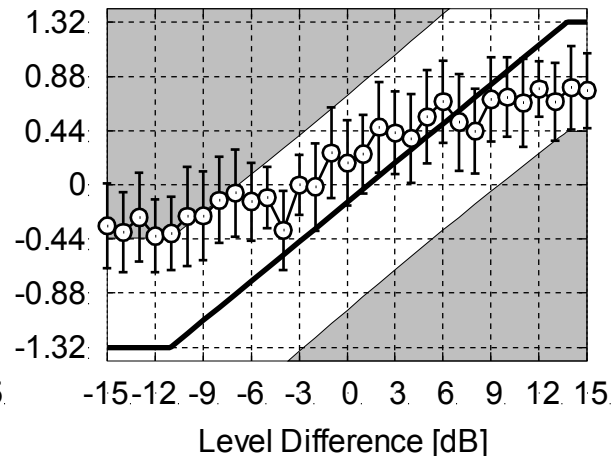
White Noise (1030ms)



Speech (1030ms)



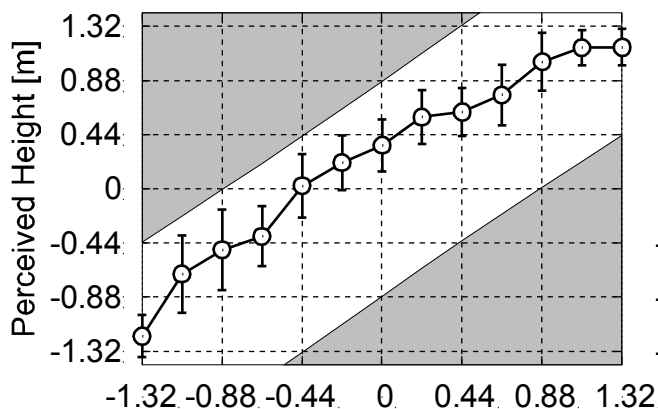
Flute (1030ms)



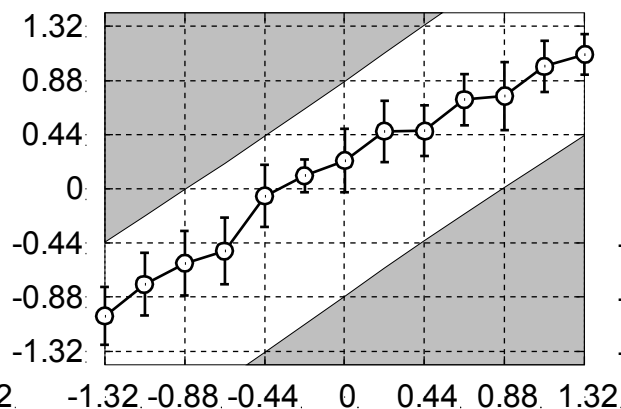
# 実験結果(統制条件)

- フルート, 残響時間1030msで弁別可能

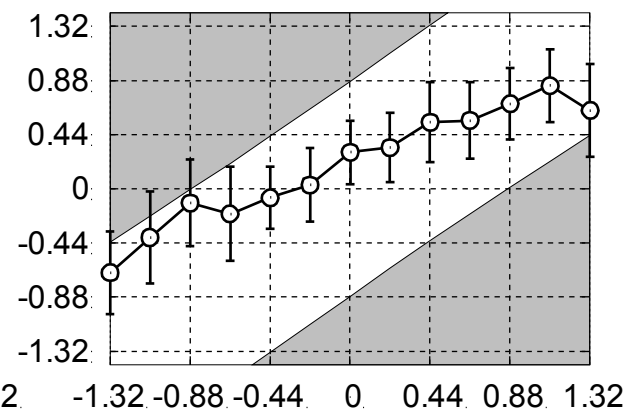
White Noise (140ms)



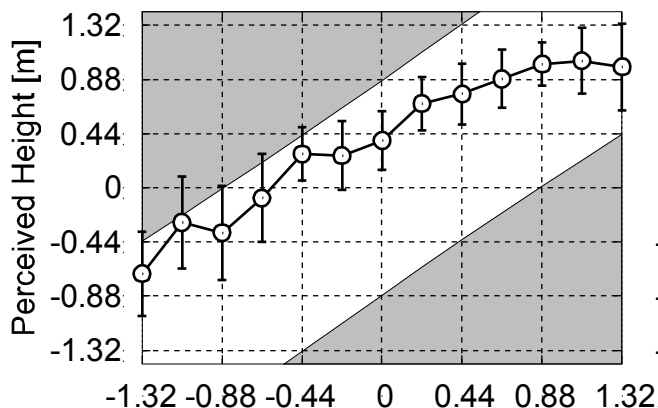
Speech (140ms)



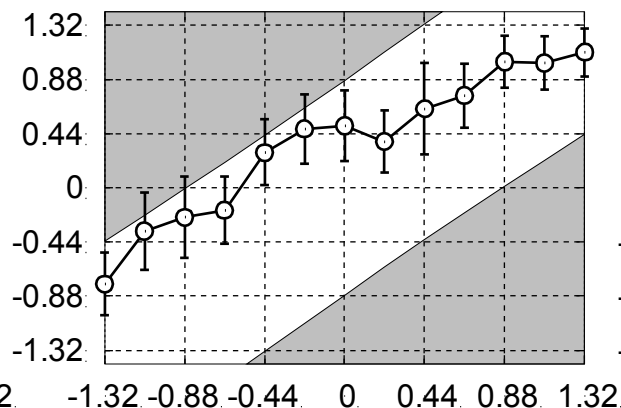
Flute (140ms)



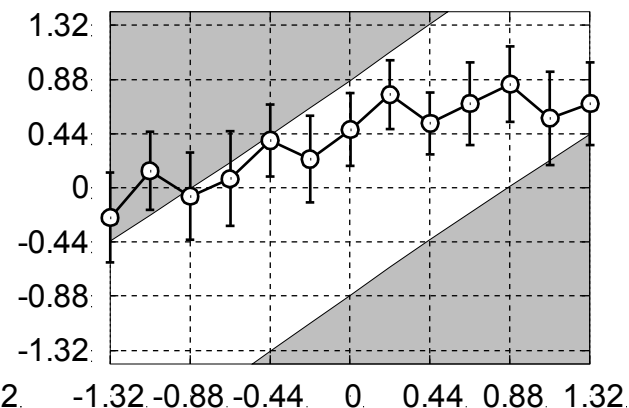
White Noise (1030ms)



Speech (1030ms)



Flute (1030ms)



Presented Height [m]

Presented Height [m]

Presented Height [m]

- パニング条件

- 5条件で高さ平均が弁別閾に入っていない
  - パニング曲線の性能は違いが分からないほど高い
- フルート, 残響時間1030ms
  - 下側において高さ平均が弁別領域に入っているの  
で, 違いを判別できる可能性あり

- 統制条件

- フルート, 残響時間1030ms
  - 下側において高さ平均が弁別領域に入っている

- 音源がフルートの場合は残響時間によって高さ知覚自体ができていない可能性あり

- 今後, 残響時間を詳細に変更した追加実験が必要

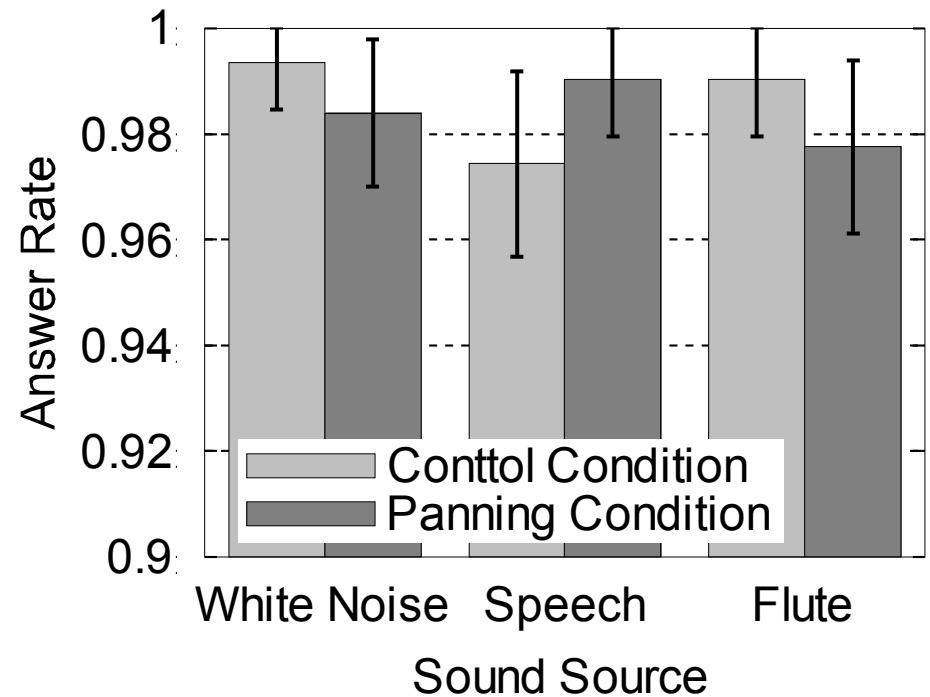
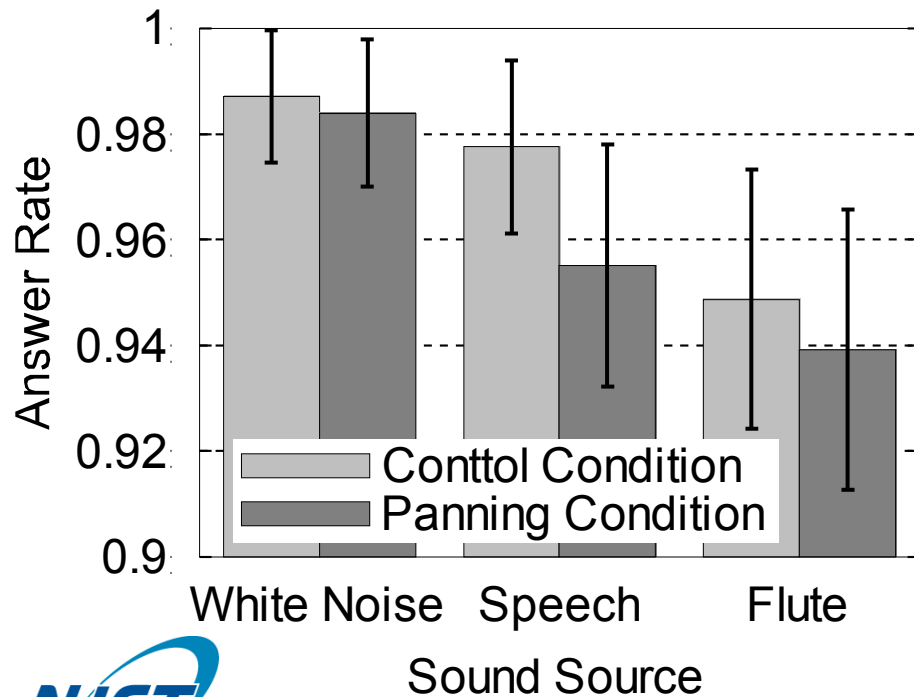
# 「音像が一つ」と回答した割合

- パニング条件

- 範囲: 音量差-3~9dB
- 音声(140ms)を除き, 統制条件と有意差なし
- 常に93.9%以上

RT 140ms

RT 1030ms



# まとめ

- 新たなる立体音響システムを提案
  - Multiple Vertical Panning (MVP)方式
  - 大画面裸眼立体映像提示システムに適合
- 聴取実験で提案システムの性能を評価
  - 垂直に27個のスピーカを配置し, 実験を実施
  - 正確な位置に合成音像の提示が可能
    - 残響時間1030 ms, 音源がフルートの場合を除く
- 今後の課題
  - 残響時間を詳細に変更して音像の高さ知覚の更なる検討
    - 音源がフルートの場合
  - 大画面裸眼立体映像提示システムと統合した視聴覚実験