

個人用三次元音場再生用 収録システムの 三次元定位の評価

木村敏幸

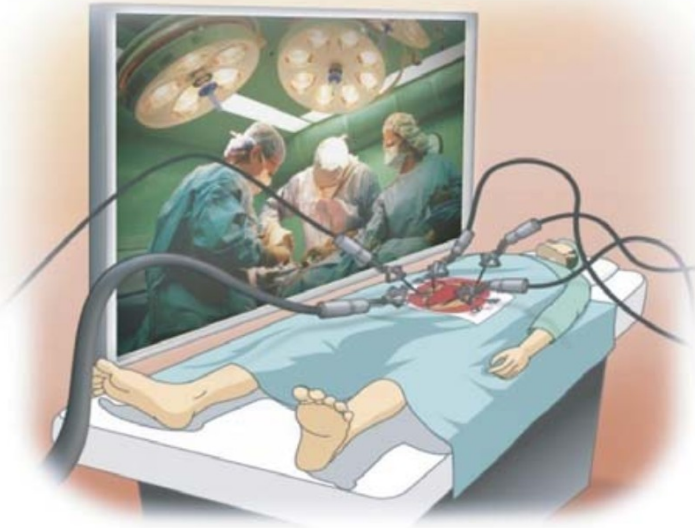
東北学院大学工学部情報基盤工学科

2022年8月4日

はじめに

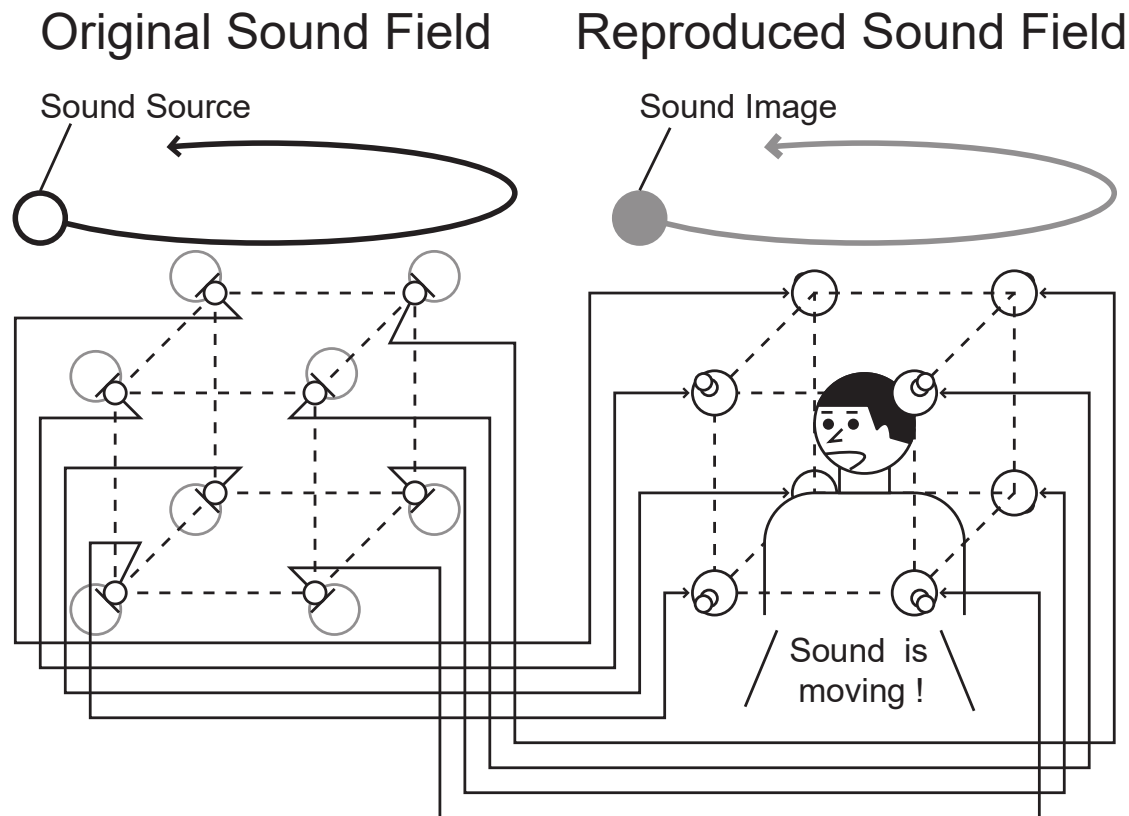
- 超臨場感コミュニケーション技術
 - 「臨場感」を体験させる技術
 - 立体映像や高臨場感オーディオ技術を使用
- 活用事例
 - 立体テレビ，遠隔医療，遠隔通信会議

出典元：榎並，岸野，電子情報通信学会誌，
Vol. 93, No. 5, pp. 363-367, May 2010.



個人用三次元音場再生

- 原音場
 - 8個の超指向性マイクロホンで音を収録
- 再生音場
 - 8個のスピーカで収録した音を再生

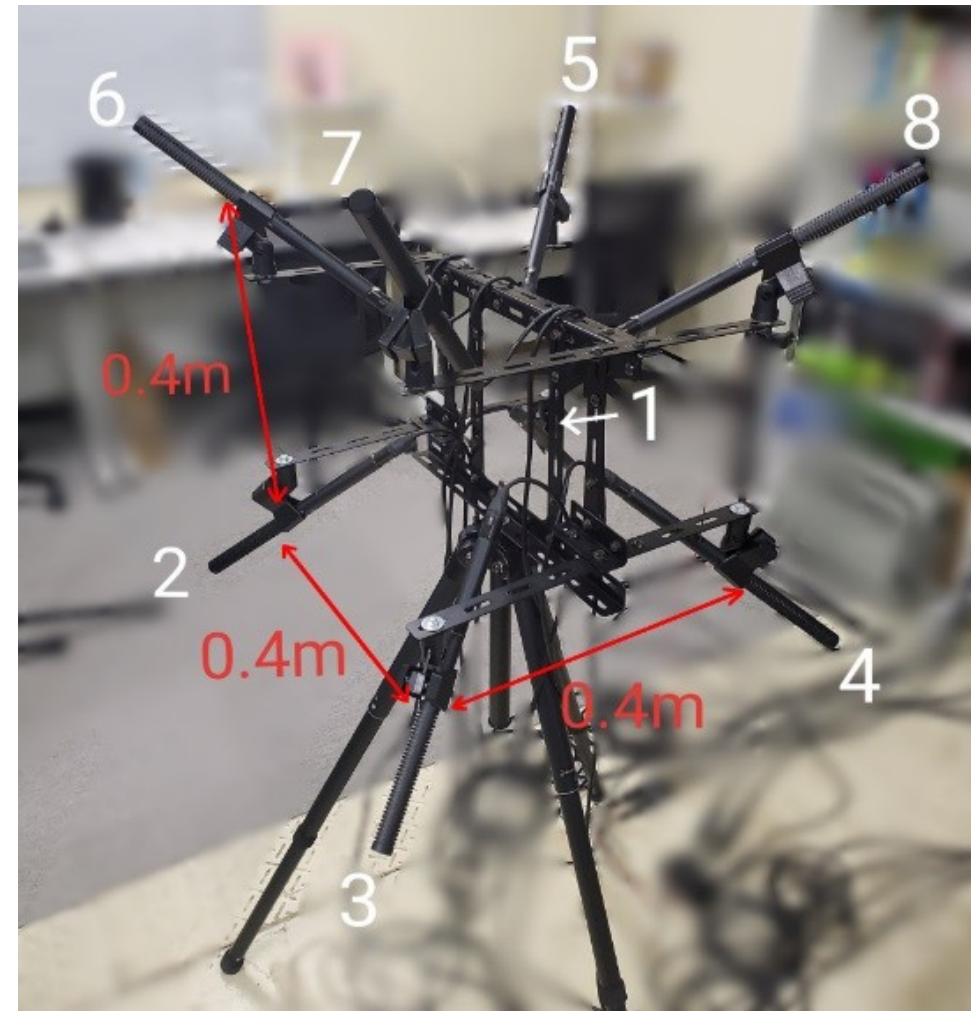


出典元：木村敏幸, "8個の指向性マイクロホンを用いた波面合成技術のコンセプトに基づいた個人用コンパクト三次元音場再生システム," 電子情報通信学会論文誌, Vol. J97-A, No. 4, pp. 284-294, April 2014.

三次元音場再生用収録システム

- マイクロホンアレイ用フレーム
 - 側方の4枚のプレートで保持
 - アレイ中心に配置しない
 - 立方体一辺：0.4 m
- 超指向性マイクロホン
 - AZDEN SGM-1000
 - 收音位置を立方体の頂点位置に配置
 - 先端から13.7 cm

出典元：羽生史, 木村敏幸, "個人用三次元音場再生システム用収録装置の主観評価," 電子情報通信学会技術研究報告, No. EA2020-81, pp. 128-133, March 2021.

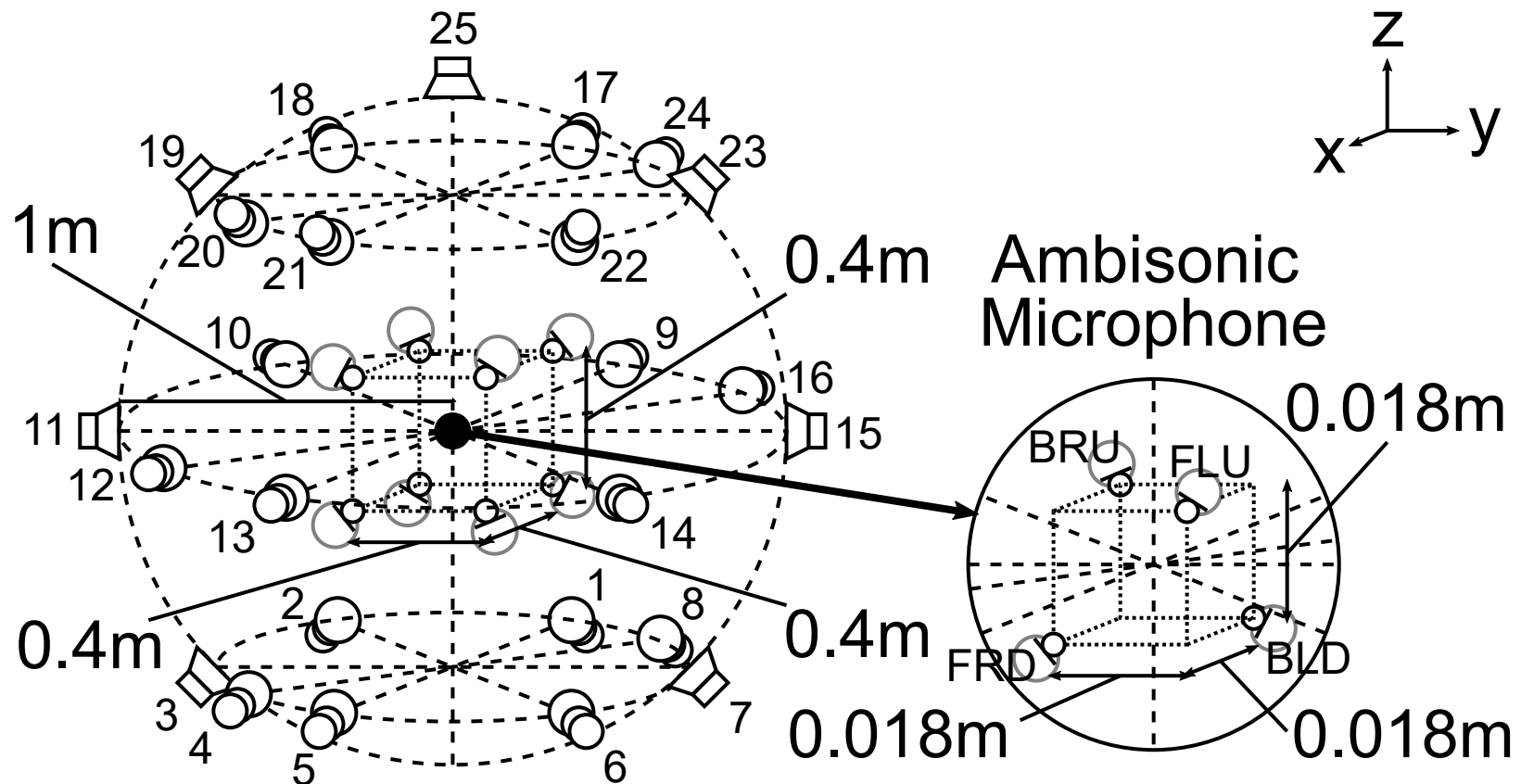


本研究の目的

- 個人用三次元音場再生
 - これまで
 - 手法を提案し，収録システムを構築
 - 二次元方向の定位性能を評価
 - 三次元方向の性能は評価せず
- 本研究の目的
 - 提案手法による三次元方向の定位性能を評価
 - インパルス応答の測定
 - 再生音の合成
 - 定位実験
 - アンビソニックマイク（従来法）と比較

インパルス応答の測定

- 実験室内で実施
 - 残響時間 : 500 ms
 - 暗騒音レベル : 32.6 dBA
 - 音圧レベル : 75.1 dBA (音源から1 m)



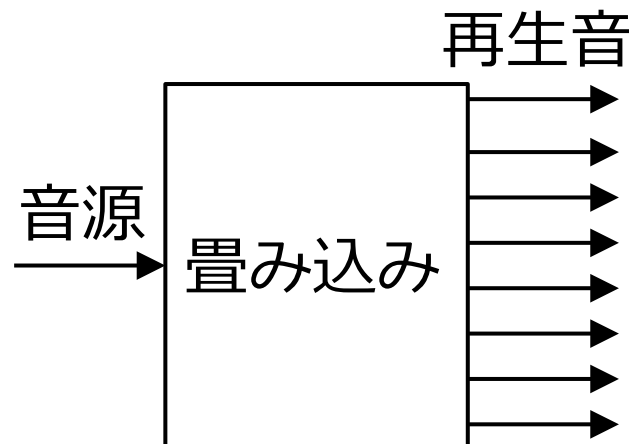
インパルス応答の測定

- アンビソニックマイク : Zoom H3-VR
- スピーカ : オーム電機 ASP-204N-K
- TSP信号
 - 標本化周波数
 - 48 kHz
 - 信号長
 - 65536点
- 同期加算
 - 9回



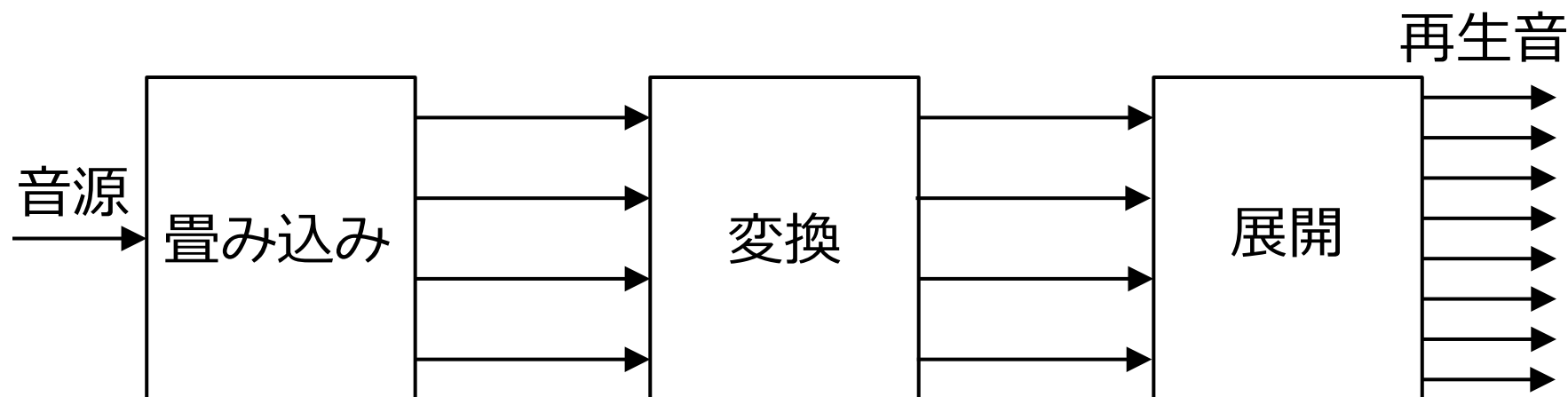
再生音の合成

- 提案する収録システム
 - 測定したインパルス応答に音源信号を畳み込み
 - 畳み込み前に残響成分を除去
 - 直接波成分以降の振幅を0に
 - 音源信号
 - 白色雑音, 音声
 - 長さ: 4秒間
 - 標本化周波数: 48 kHz



再生音の合成

- アンビソニックマイク
 - 測定したインパルス応答に音源信号を畳み込み
 - 畳み込み前に残響成分を除去
 - 直接波成分以降の振幅を0に
 - 畳み込み後, フォーマットを変換
 - W (無指向性), X, Y, Z (双指向性)
 - 変換後, 8チャンネルに展開



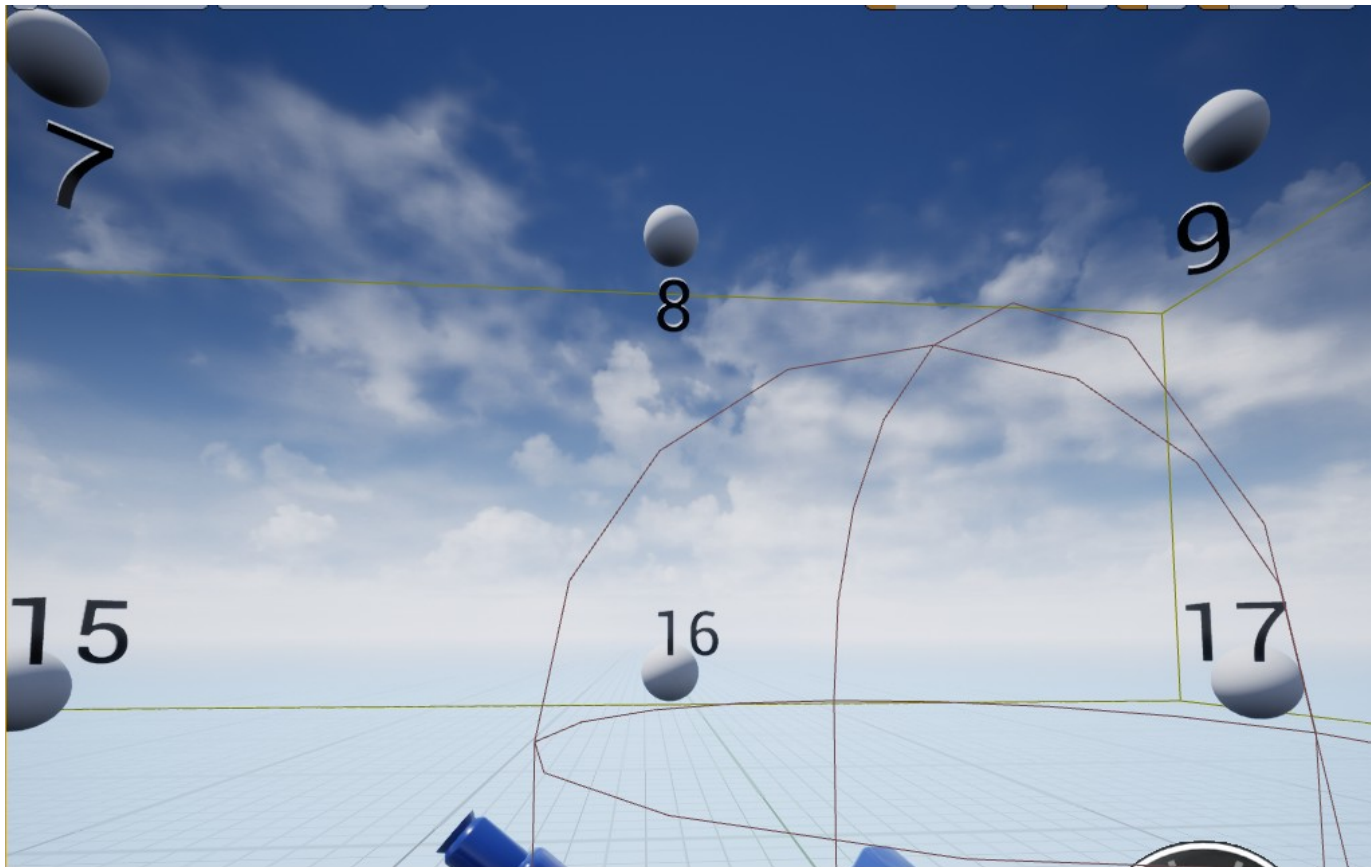
定位実験

- 実験室内の一角
 - 実験参加者
 - スピーカアレイ内に座り
ヘッドマウント
ディスプレイを装着
 - 頭部中心の高さ
 - 地面から1.4 m
 - アレイ中心と同じ
 - 音圧レベル
 - 約70 dBA : アレイ中心



定位実験

- ヘッドマウントディスプレイ
 - Microsoft : Hololens2
 - 三次元仮想空間映像を常時提示
 - 番号付きオブジェクトを25個配置

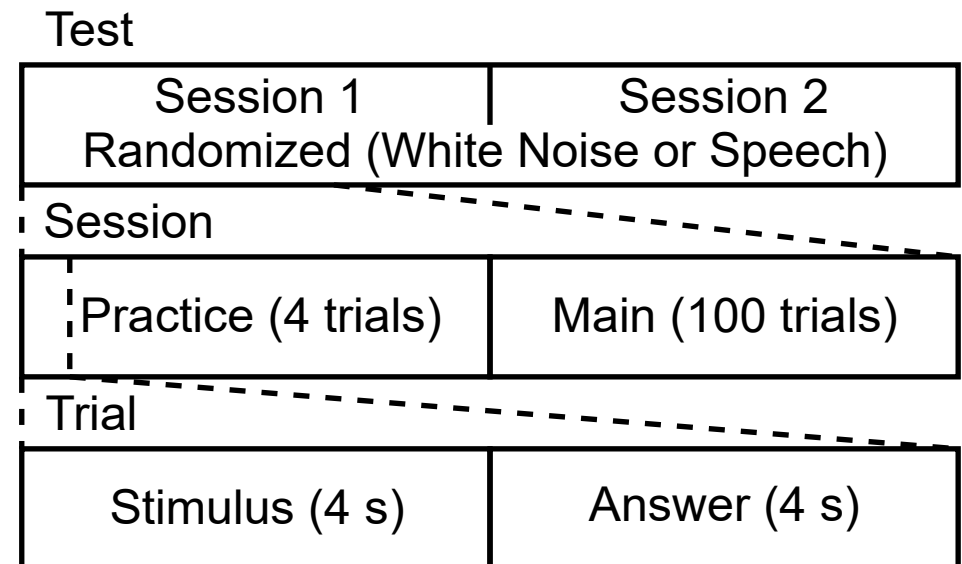


定位実験

- 実験参加者
 - 10名
 - 実験開始前
 - 実験内容に関する説明を受ける
 - 自分の意思で実験に参加できる
 - 個人情報 は保護される
 - 説明後
 - 参加に同意したら, 同意書に署名

定位実験

- セッション順序
 - 実験参加者ごとにランダムイズ
- 本試行の内訳
 - 25 (方向) ×2 (収録装置) ×2 (繰り返し)
- 提示順序
 - 実験参加者ごとにランダムイズ
- 試行
 - 定位方向の番号を回答
 - 頭部は自由に回転可能



実験結果

- 知覚方向誤差

- 方位角, 仰角の誤差で検討するのは不適切

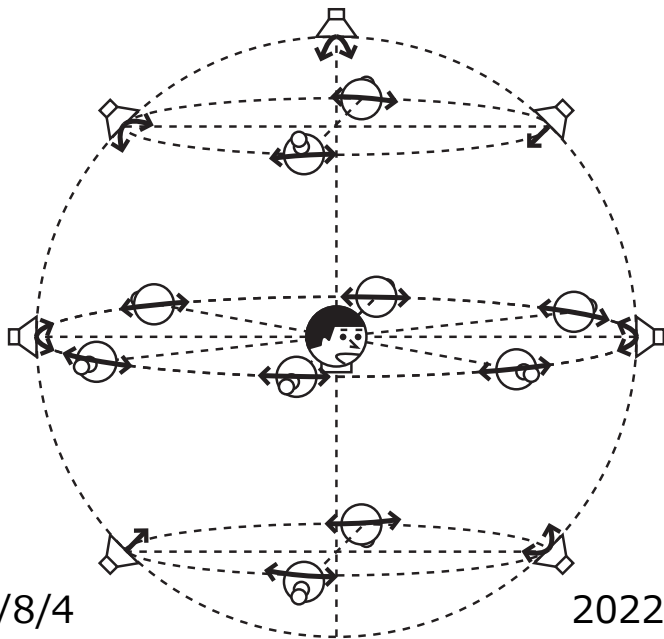
- 真上方向 (仰角90度) の場合, 誤差方向が分からない

⇒ 方位角及び仰角を「水平角」と「垂直角」に変換

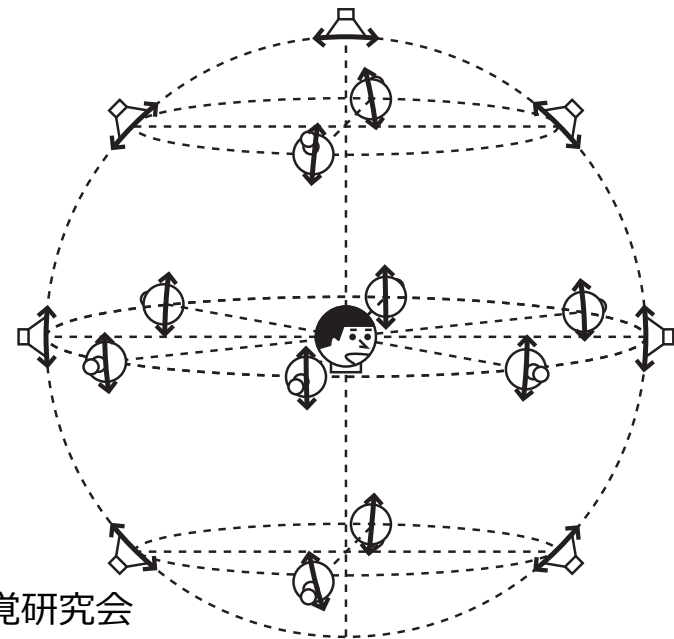
- 水平角 : 聴取者が提示方向を見たときの左右方向の変位

- 垂直角 : 聴取者が提示方向を見たときの上下方向の変位

水平角



垂直角



実験結果

- 知覚方向誤差

$$\theta' = \tan^{-1} \frac{y'}{x'}, \phi' = \sin^{-1} \frac{z'}{\sqrt{x'^2 + y'^2 + z'^2}}$$

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathbf{R}_y(-\phi_0) \mathbf{R}_z(-\theta_0) \begin{pmatrix} \cos \theta \cos \phi \\ \cos \theta \sin \phi \\ \sin \phi \end{pmatrix}$$

$$\mathbf{R}_y(-\phi_0) = \begin{pmatrix} \cos \phi_0 & 0 & \sin \phi_0 \\ 0 & 1 & 0 \\ -\sin \phi_0 & 0 & \cos \phi_0 \end{pmatrix} \mathbf{R}_z(-\theta_0) = \begin{pmatrix} \cos \theta_0 & \sin \theta_0 & 0 \\ -\sin \theta_0 & \cos \theta_0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- θ', ϕ' : 変換した水平角, 垂直角

- θ, ϕ : 回答方向の方位角, 仰角

- θ_0, ϕ_0 : 提示方向の方位角, 仰角

- $\mathbf{R}_z(-\theta_0), \mathbf{R}_y(-\phi_0)$: 水平方向, 上下方向の回転操作

実験結果

- 知覚方向誤差

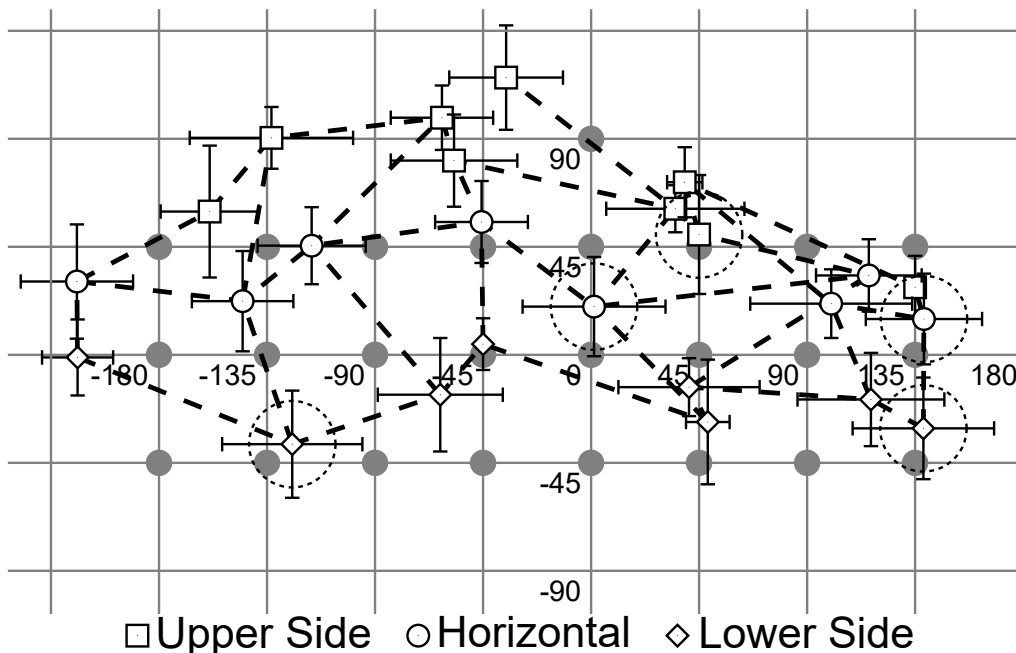
- アンビソニックマイク

- 正しい定位を示す提示方向の回答が少ない

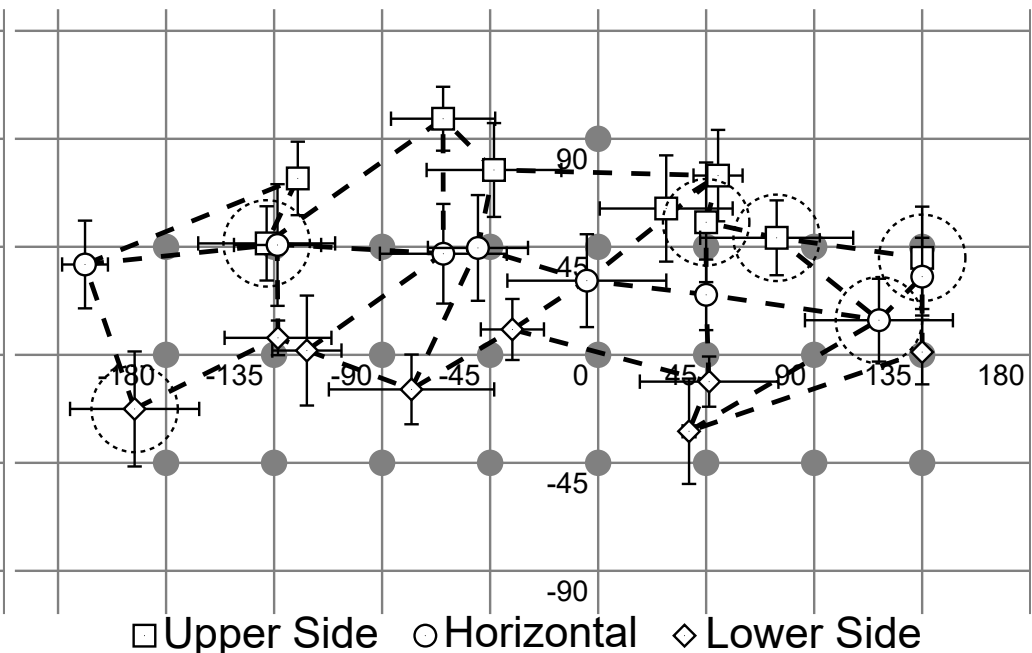
- 白色雑音 : 5, 音声 : 6

⇒ **良い定位性能とは言えない**

白色雑音



音声



実験結果

- 知覚方向誤差

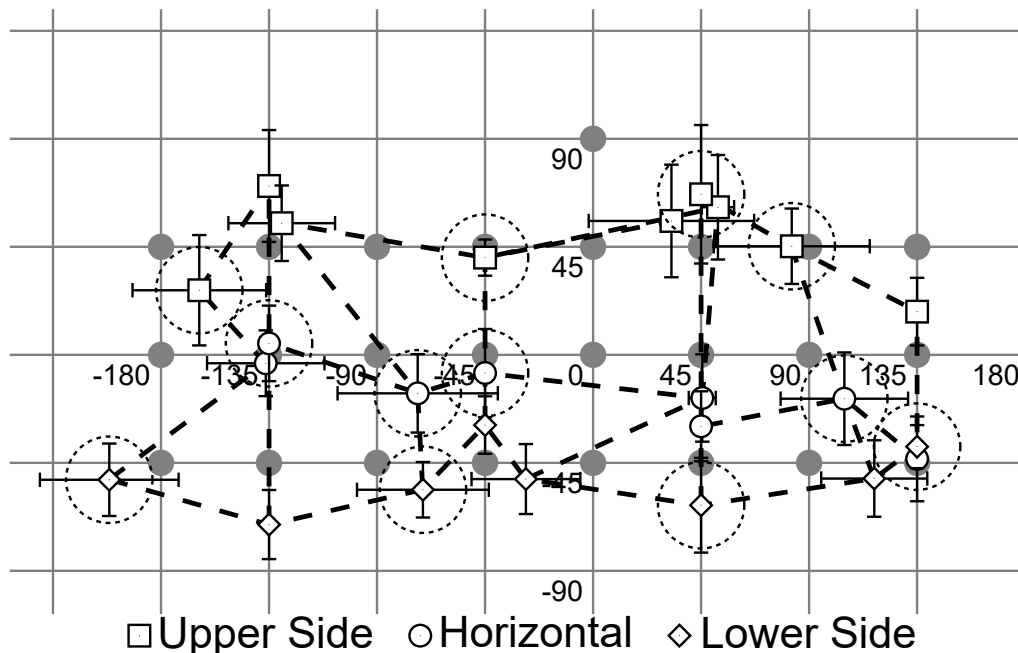
- 提案する収録システム

- 正しい定位を示す提示方向の回答が増加

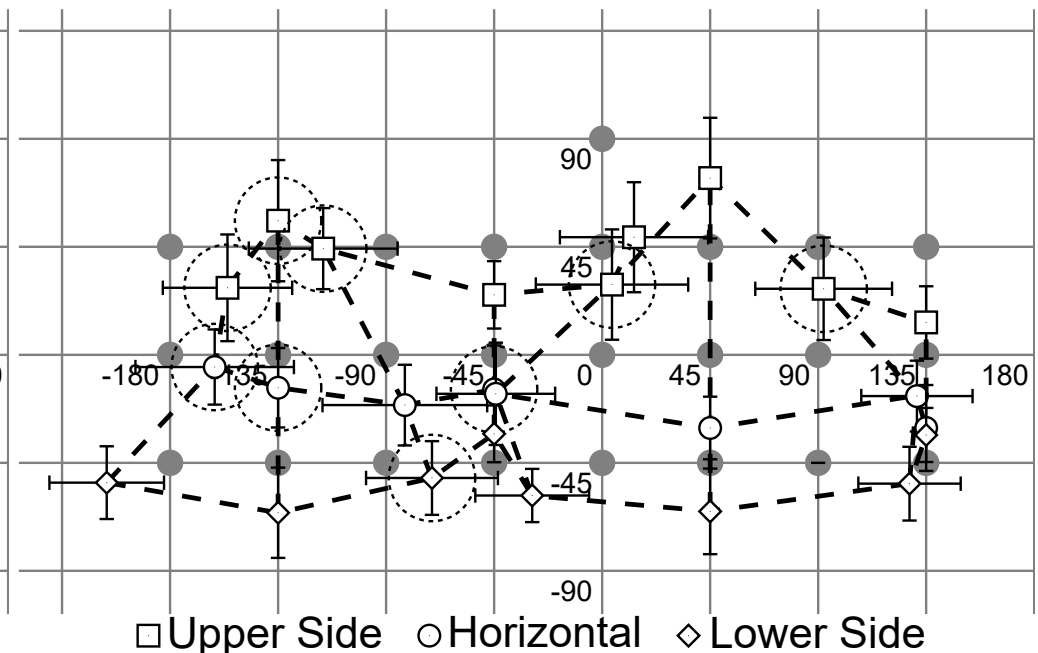
- 白色雑音 : 12, 音声 : 9

⇒ **定位性能はアンビソニックマイクより良い**

白色雑音



音声



実験結果

- 正答率

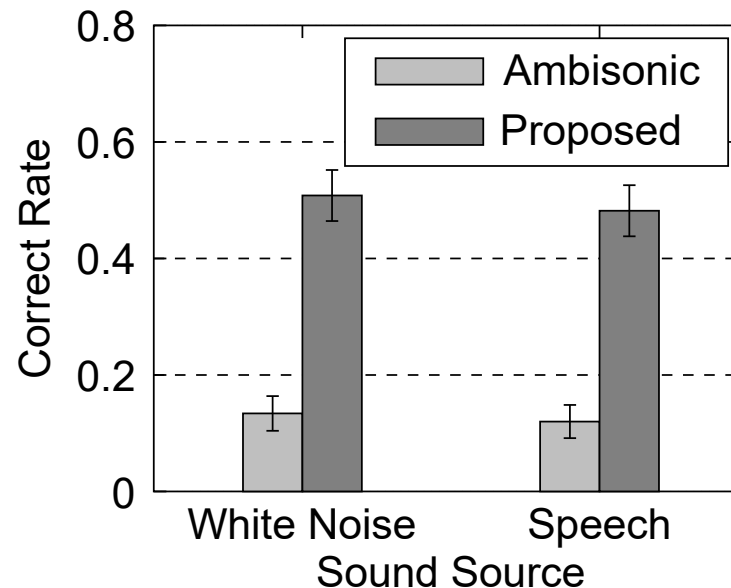
- 回答数：500

- 25（方向）×2（繰り返し）×10（参加者数）

- 提案する収録システムの方が高い

- Fisherの正確確率検定で0.1%水準有意

⇒ **定位性能はアンビソニックマイクより良い**



まとめ

- 個人用三次元音場再生用収録システムの評価
 - 三次元方向の定位性能を評価
 - 定位性能をアンビソニックマイクと比較
 - ⇒ アンビソニックマイクよりも良い定位性能
- 今後の課題
 - 映像と組み合わせた遠隔操作システムの構築
 - 視聴覚提示システムを用いた検討
 - 三次元音場再生の必要条件の緩和の可能性の検討