

機械学習とアレイ信号処理を 用いたロボット聴覚に関する研究

ロボット・制御科 研究員 清野若菜

質問はメールにて事務局までお気軽にお問い合わせください。
問い合わせ先：福島県ハイテクプラザ 企画連携部産学連携科
e-mail : hightech-renkei@pref.fukushima.lg.jp

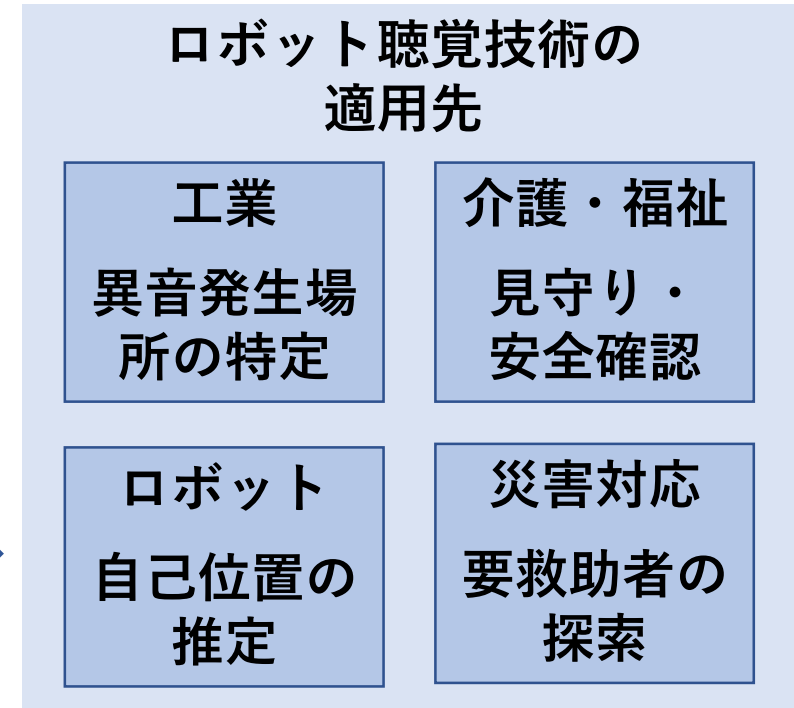
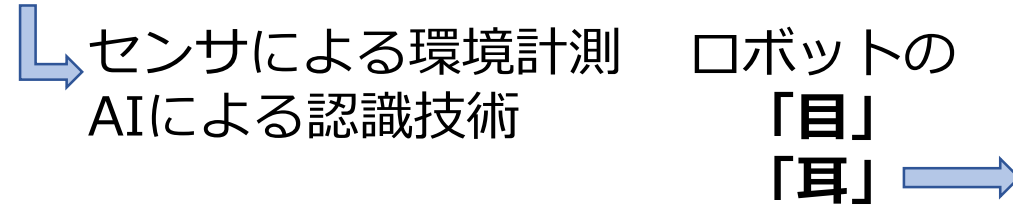
目次

1. 背景と目的
2. 小型音源定位システム
 - ・ マイクアレイの試作
 - ・ 実験
 - ・ 結果及び考察
3. ロボット聴覚システム
 - ・ データセット
 - ・ 学習モデル
 - ・ 水流音の位置推定及び検出
4. まとめ

1. 背景と目的

背景：

2050年までに、**AIロボット**が広く浸透



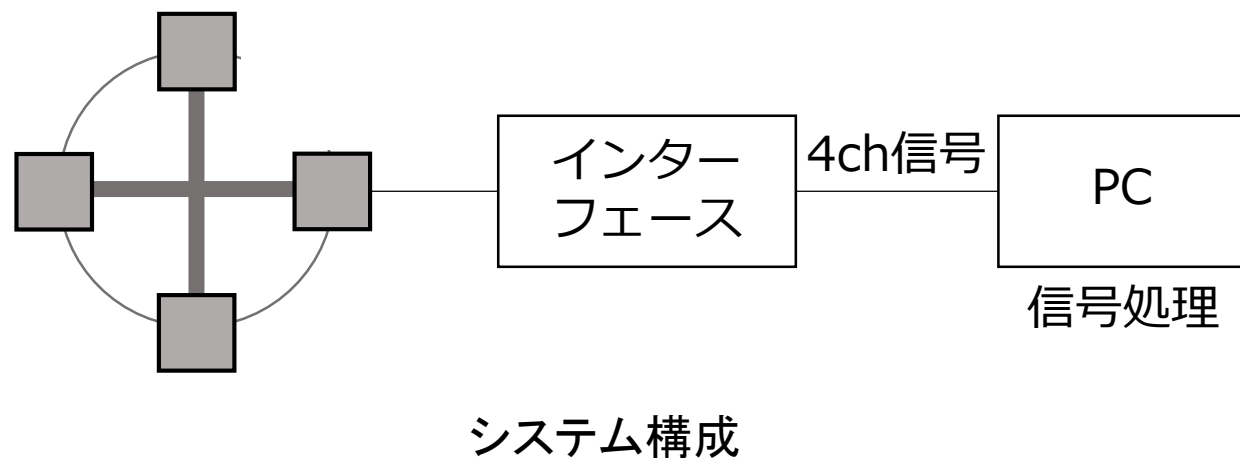
目的：

- ① 小型AIサービスロボットにも搭載可能な少数チャンネルマイクロホンアレイの試作及び性能評価
- ② 機械学習を用いた環境中の音の「方向」と「対象」を検出するロボット聴覚システムの実装及び性能評価

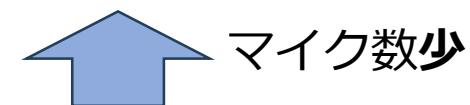
2. 小型音源定位システム

(1) 4chマイクロホンアレイの試作

- マイク：システムインフロンティア CSMIC-S1R1
- インターフェース：システムインフロンティア RASP-ZX
- アレイ半径：36.5 mm



【試作品】4chマイクロホンアレイ



8chアレイ



25ch球形アレイ

2. 小型音源定位システム

(2) 音源定位実験

- 場所：メカトロ研究室
- マイクアレイ：4ch, 8ch(システムインフォニア TAMAGO)
- 音源：配管スチーム音

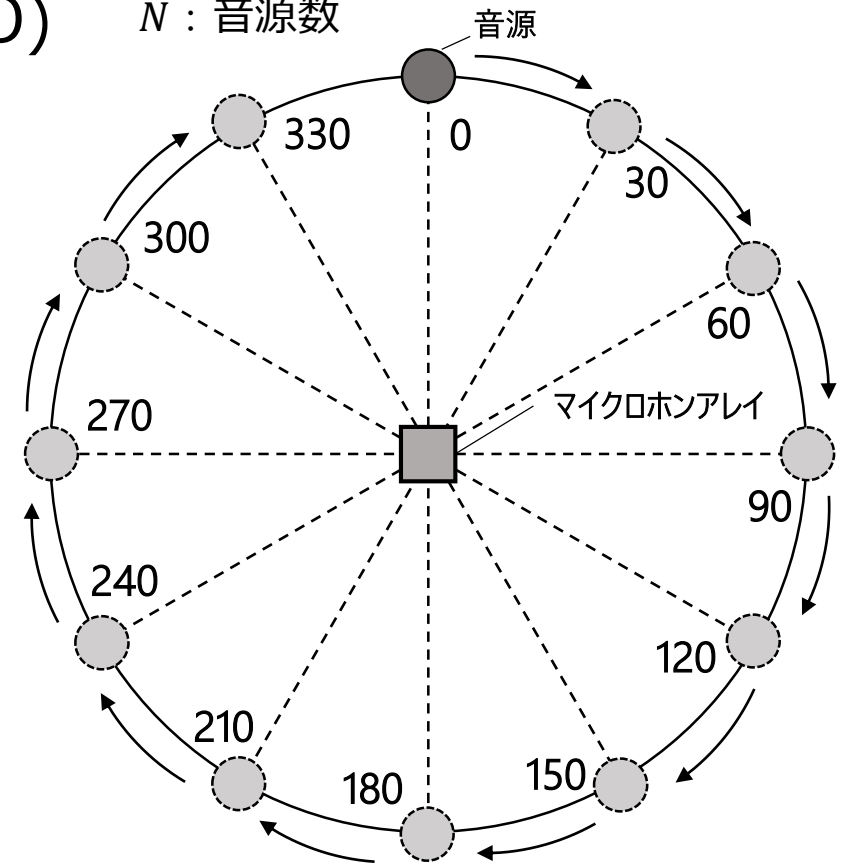
MUSIC法による空間スペクトルの算出方法

$$P(\theta) = \frac{|\mathbf{a}^H(\theta)\mathbf{a}(\theta)|}{\sum_{i=N+1}^M |\mathbf{a}^H(\theta)\mathbf{e}_i|^2}$$

$P(\theta)$: 空間スペクトル
 $\mathbf{a}(\theta)$: アレイ・マニフォールド・ベクトル
 \mathbf{e}_i : 固有値番号*i*番目の固有ベクトル
 M : マイクロホン数
 N : 音源数

表 音源定位の実装条件

パラメータ	値
MUSICアルゴリズム	一般化固有値分解(SEVD)
STFT点数	512ポイント
サンプリング周波数	16000[Hz]
フレーム長	32[ms] (512ポイント)
フレームシフト	10[ms] (160ポイント)
使用周波数	500~6400[Hz]
アレイ・マニフォールド・ベクトルの分解能	1[°]



実験配置図

2. 小型音源定位システム

(3) 音源定位実験 結果

【平均定位誤差】 4ch : 4.4[°] 8ch : 2.8[°]

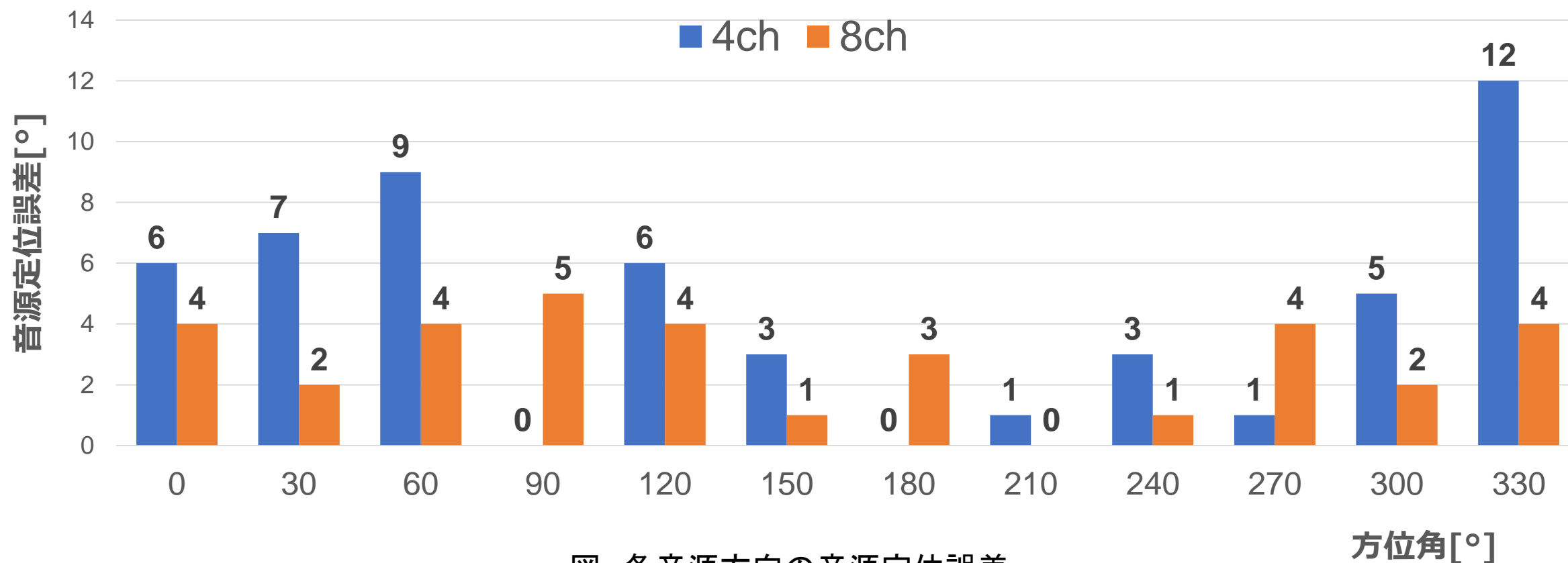


図 各音源方向の音源定位誤差

3. ロボット聴覚システム

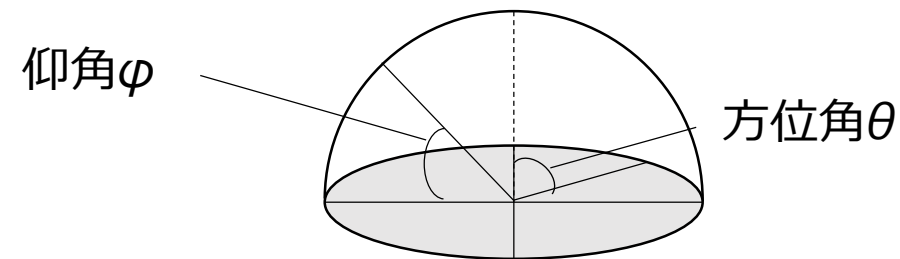
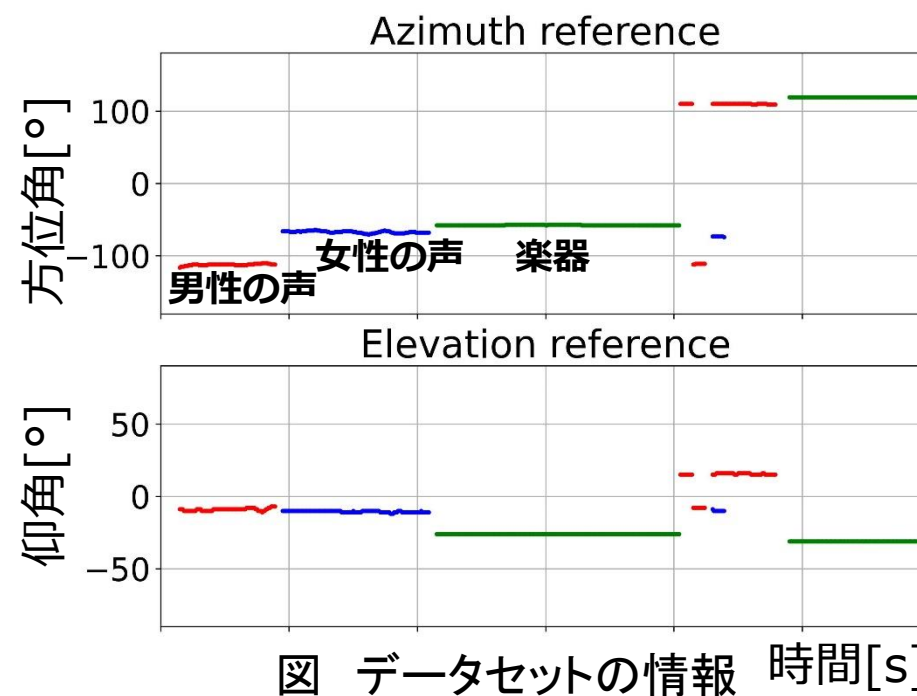


表 データセットのクラス分類

ID	環境音の種類	6	歩く音、足音
0	女性の話し声	7	ドアの開閉
1	男性の話し声	8	音楽
2	拍手	9	楽器
3	電話	10	水道、蛇口
4	笑い声	11	ベル
5	家庭内の音	12	ノック



3. ロボット聴覚システム

(2) 学習モデル

- DCASE2023 Task3 「音響イベントの位置推定と検出」
ベースラインモデルを使用 (CRNNベース)
- 学習エポック数：200

うまく推論できている例

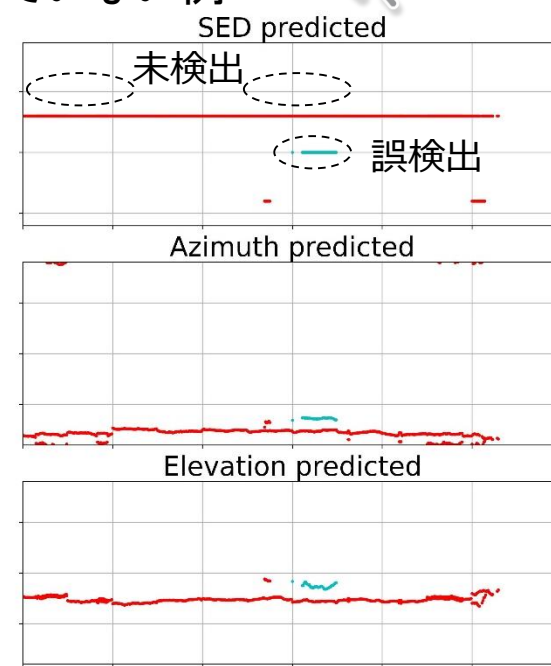
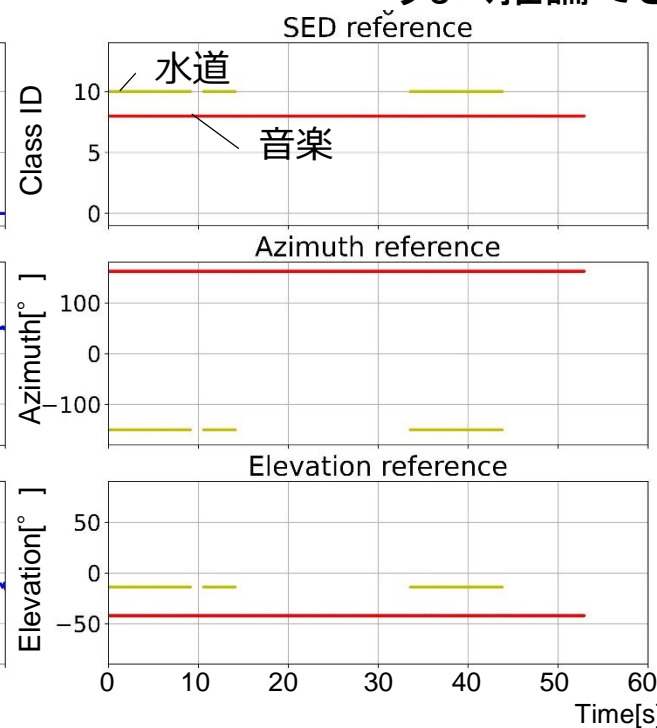
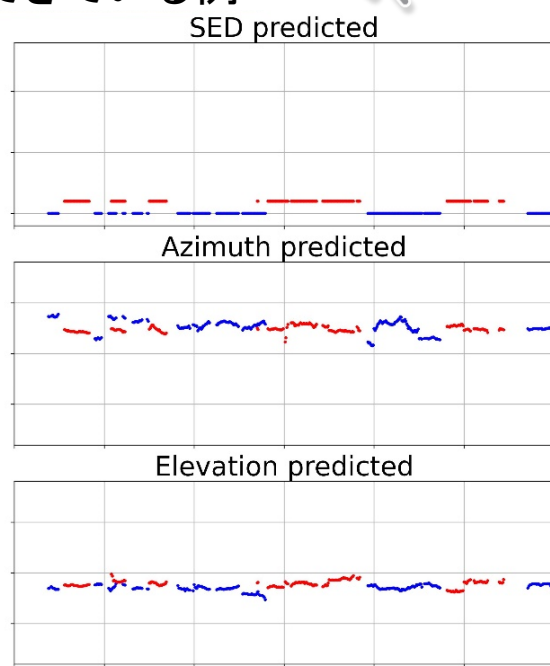
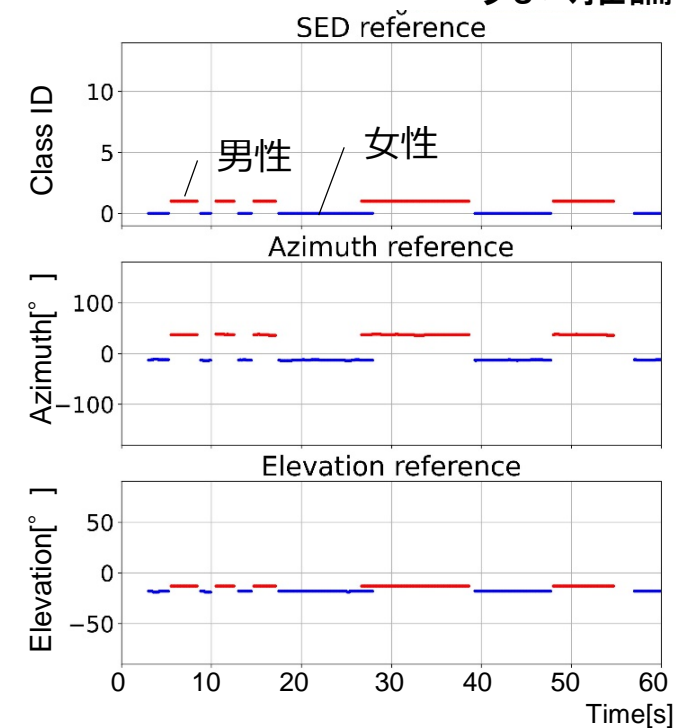


うまく推論できていない例



表 【結果】ベースラインモデルの性能評価

性能評価指標	値 ()内は公式記録
Error rate	0.75 (0.57)
Localization error[°]	65.8 (21.6)



3. ロボット聴覚システム

(3) 水流音の位置推定及び検出

- 前述のベースラインモデルをファインチューニングし、独自の水流音データを追加学習
- 音源：水道の水流音（ICレコーダーで録音）
- 形式：24kHz 16bit、1次アンビソニックス(4ch)
- 信号長：60[s]
- データ数：学習用12, テスト用7
- 学習エポック数：50

表 水流音の推論結果

性能評価指標	値
Error rate	0.67
Localization error[°]	59.5

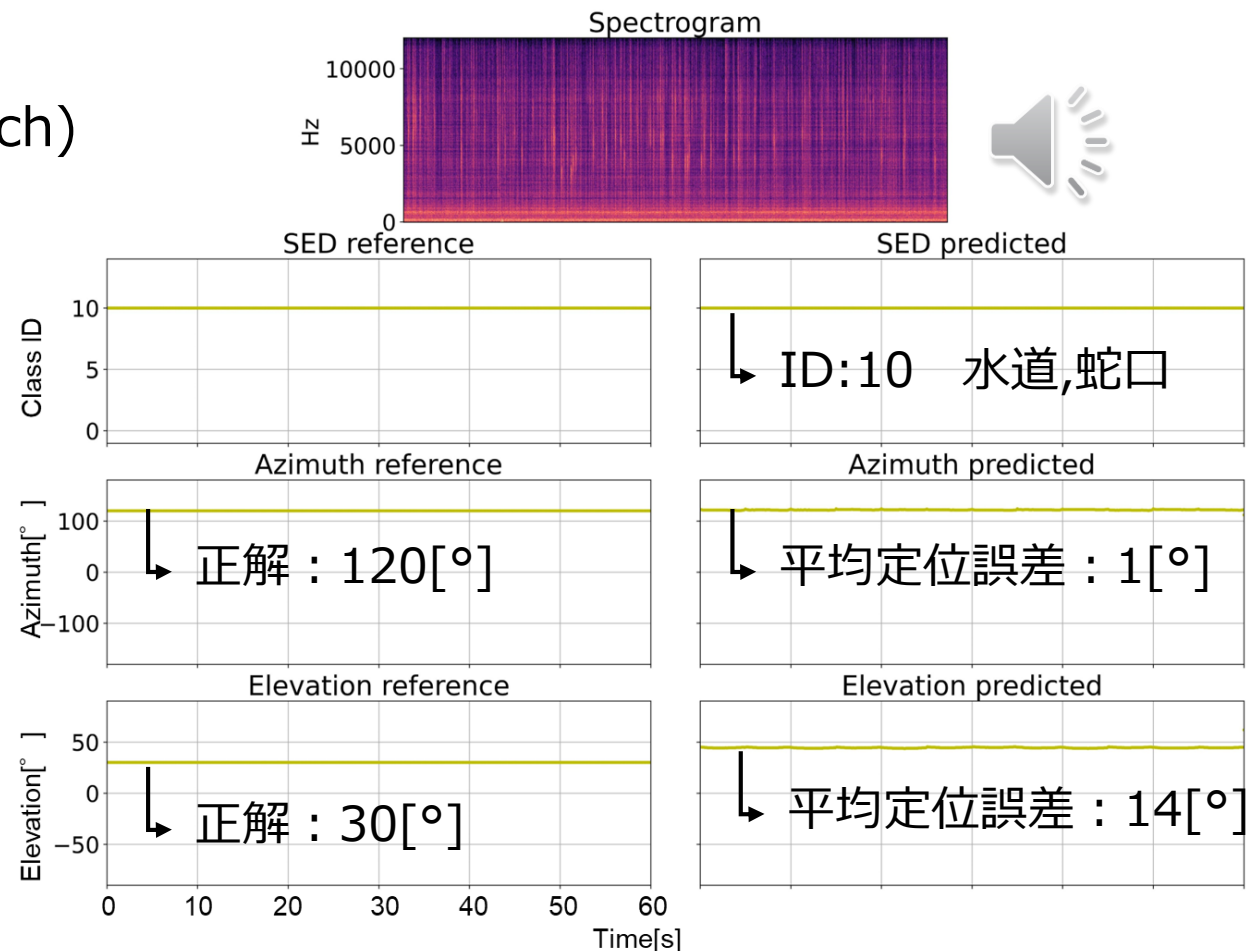
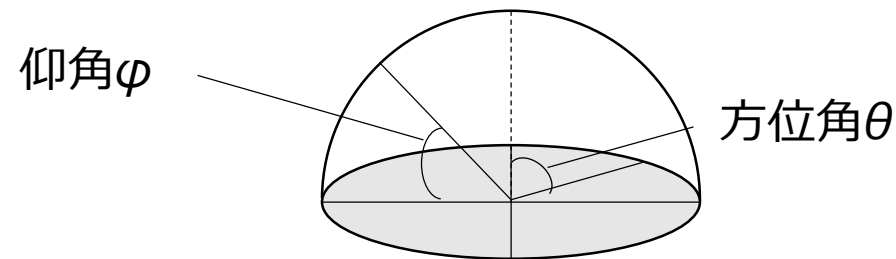


図 推論結果の可視化

4. まとめ

1. アレイ信号処理を用いた少数チャンネルマイクロホンアレイによる音源定位性能の評価及び、機械学習を用いた環境中の音の「方向」と「対象」を検出するシステムの実装と評価を行った。
2. 小型音源定位システム
 - ・ 水平方向 1 音源を対象とした音源定位では、4chのマイクロホンアレイでも8chと同等の精度であることがわかった。
3. ロボット聴覚システム
 - ・ 既存の学習済みモデルを用いて独自の検出対象音を追加学習することで、各現場に使用できるシステムの構築を省力化できる可能性があることがわかった。