

RoboCup@Homeの ヒューマンインタラクションタスク に向けた解法の提案

九州工業大学 大学院 生命体工学研究科
人間知能システム工学専攻
人間知能機械講座 田向研究室

矢野 優雅, 松本 生弥, 福田 有輝也, 小野 智寛, 田向 権

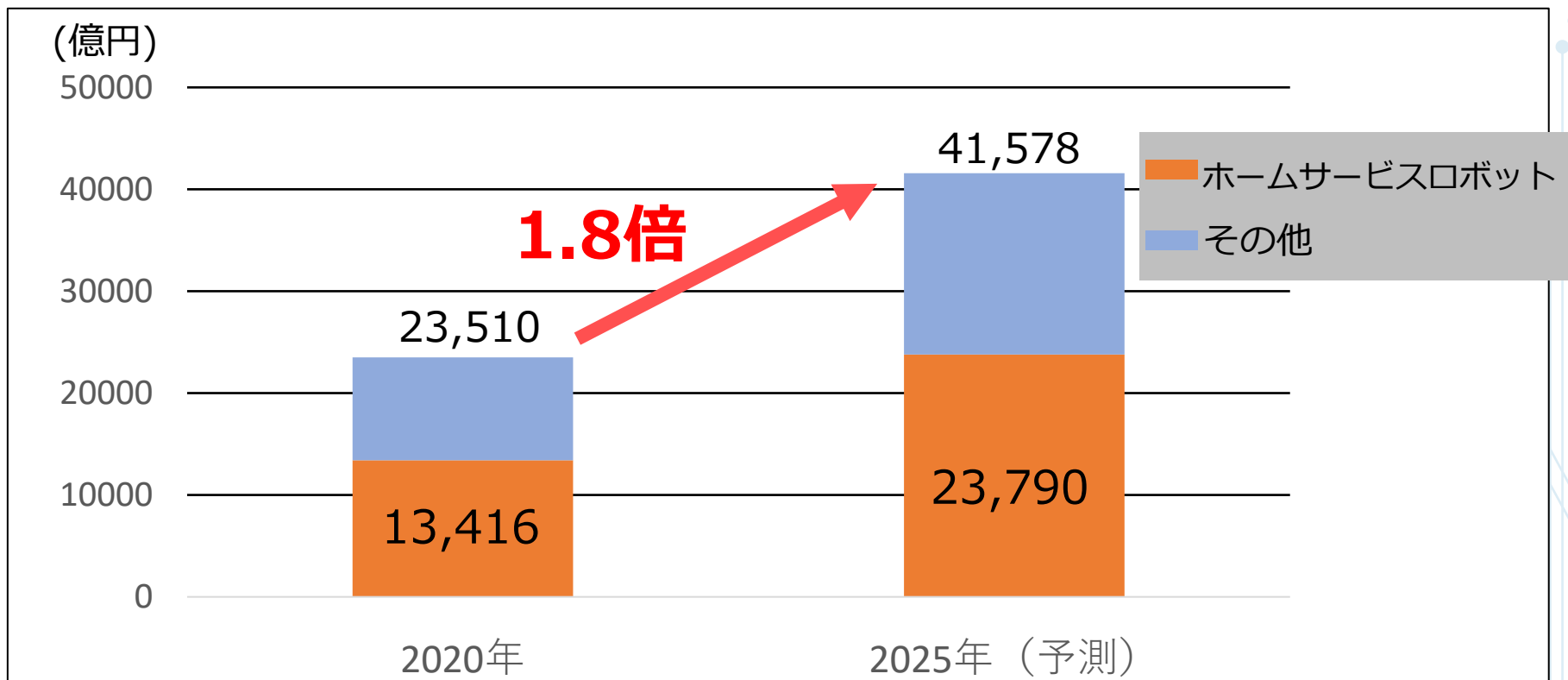


研究背景：サービスロボットの需要拡大

- ✓ 少子高齢化やコロナの影響で需要が拡大
- ✓ 家庭で動く人間と共存可能なサービスロボットが求められる

➡ ヒューマンインタラクションが重要

サービスロボットの市場規模と今後の展望^[1]



研究背景：RoboCup@Home^[2]

- ホームサービスロボットの発展を目的とする国際的な競技会
- TOYOTA社が開発したHSR[†]^[3]を用いて
家事やヒューマンインタラクションなどのタスクを行う



RoboCup@Home

本研究はヒューマンインタラクション重視のタスクに注目する

[2] RoboCup@Home, <http://www.robocupathome.org/>, (accessed 2022-10-24).

[†] Human Support Robot [3] T. Yamamoto, et al., ROBOMECH Journal, 2019.

Find My Mates (FMM)

家に訪れている4人のゲスト
を見つけホストに報告するタスク



タスク内容	必要な技術	回数(人)	合計点
メインゴール & ボーナスゴール			
ゲストの位置を報告	人物認識, 発話	3	350
位置を明示的に示す	画像による報告	3	100
ゲストの特徴を報告する	属性推定, 音声認識	3	550
減点項目			
ゲストからロボットに近づく	ナビゲーション	2	-150
合計			1000

ヒューマンインタラクションの必須要素を一度に評価できる

研究目的・目標

研究目的

人間と協調動作するホームサービスロボットの実現

研究目標

- 以下のシステムをHSRで実現する
 1. 3次元位置を含む人物認識
 2. 属性推定手法を用いた, 人物の特徴抽出
 3. 音声認識
- FMMで満点を獲得する

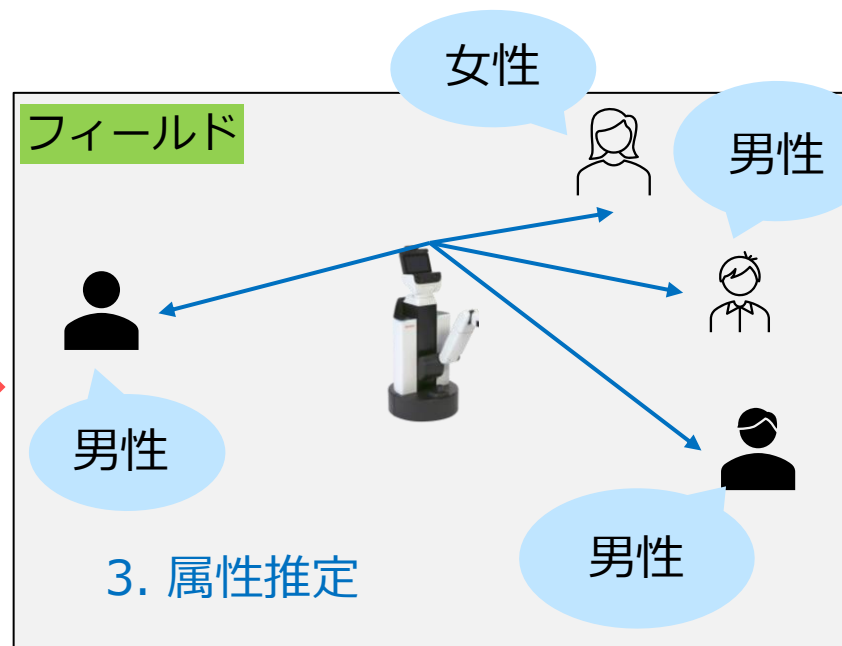
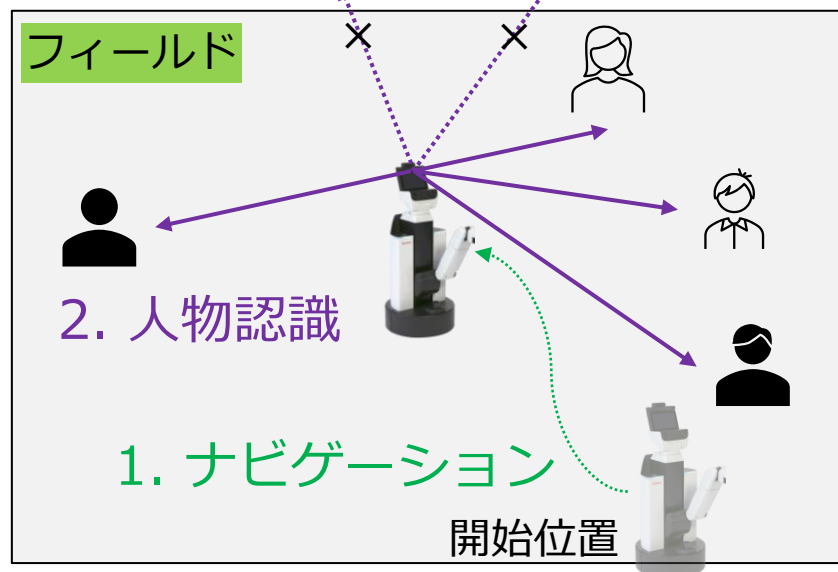
提案手法の概要 (1/2)

解法

画像を用いて，全てのゲストの
位置情報や特徴をホストに報告する

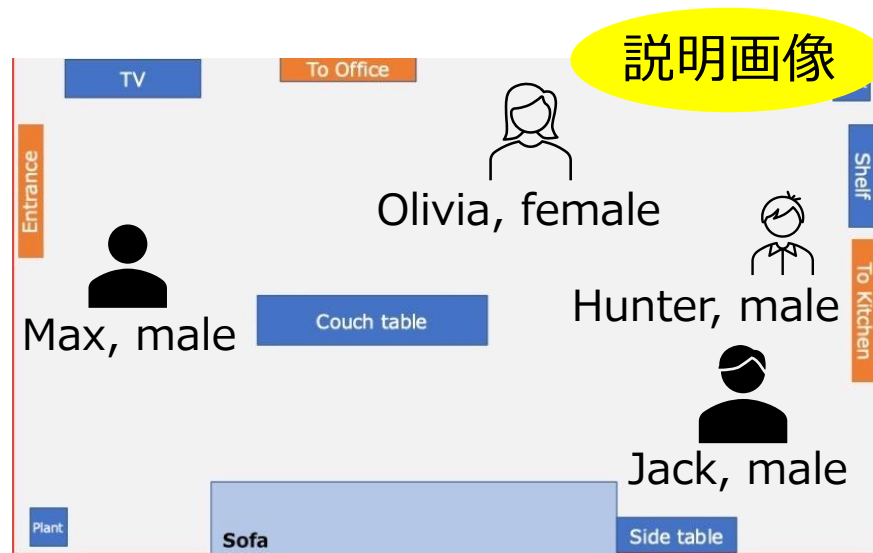
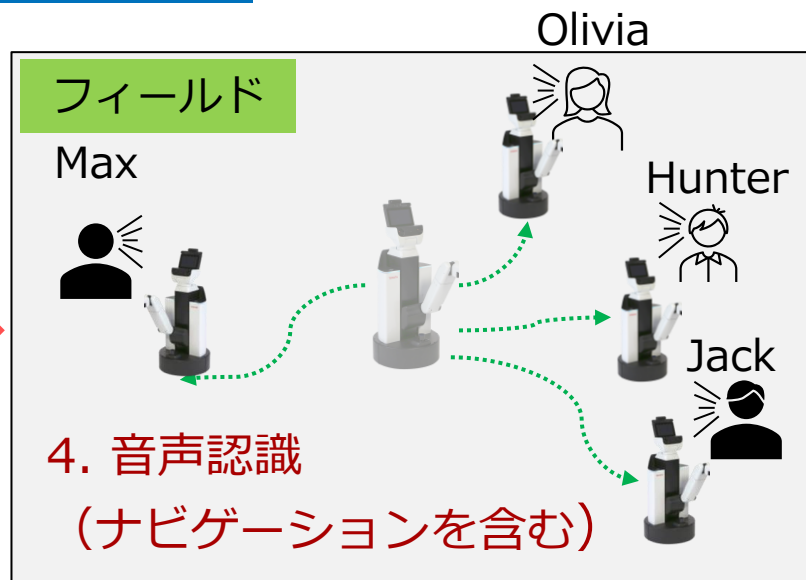
提案手法

フィールド外



提案手法の概要 (2/2)

提案手法



生成した画像（説明画像）により、
ゲストの画像・位置・名前・性別を同時に報告する

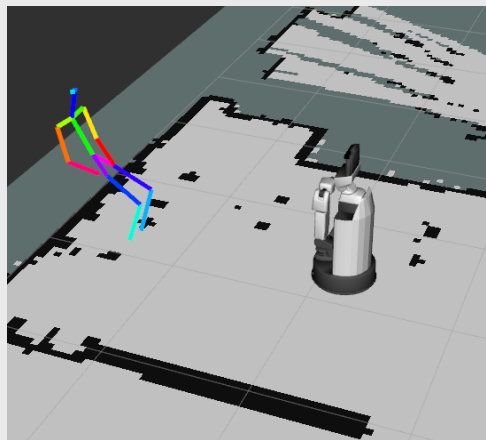
満点を獲得できる

3次元位置を含む人物認識

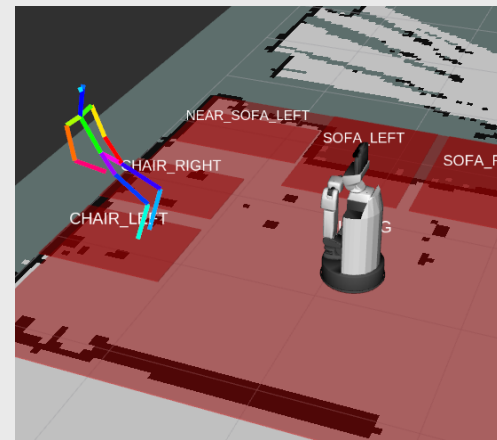
- 人物認識手法：Light-weight Human Pose Estimation^[4]
 1. HSRより取得したRGB画像を用いて人物認識を行う
 2. Depth画像も用いて、3次元での位置を算出
 3. 説明画像作成のために、どの椅子に座っているかも合わせて算出



人物認識



3次元地図上にマッピング



家具情報の意味づけ

左側の椅子に座っていることを判定可能

属性推定

- 属性推定手法：Class-Specific Residual Attention^[5]

- ✓ 性別を含む**14項目の属性を抽出**できる
- ✓ 人物認識と組み合わせて属性推定を行う
- ✓ FMMでは、性別のみを利用している



人物認識を用いた人物画像抽出

項目	出力
性別	男
服	Tシャツ
ズボン	ジーンズ
マスク	着用
髪の毛の長さ	短い

属性推定結果（一部）

RoboCup@Homeにおける音声認識の課題

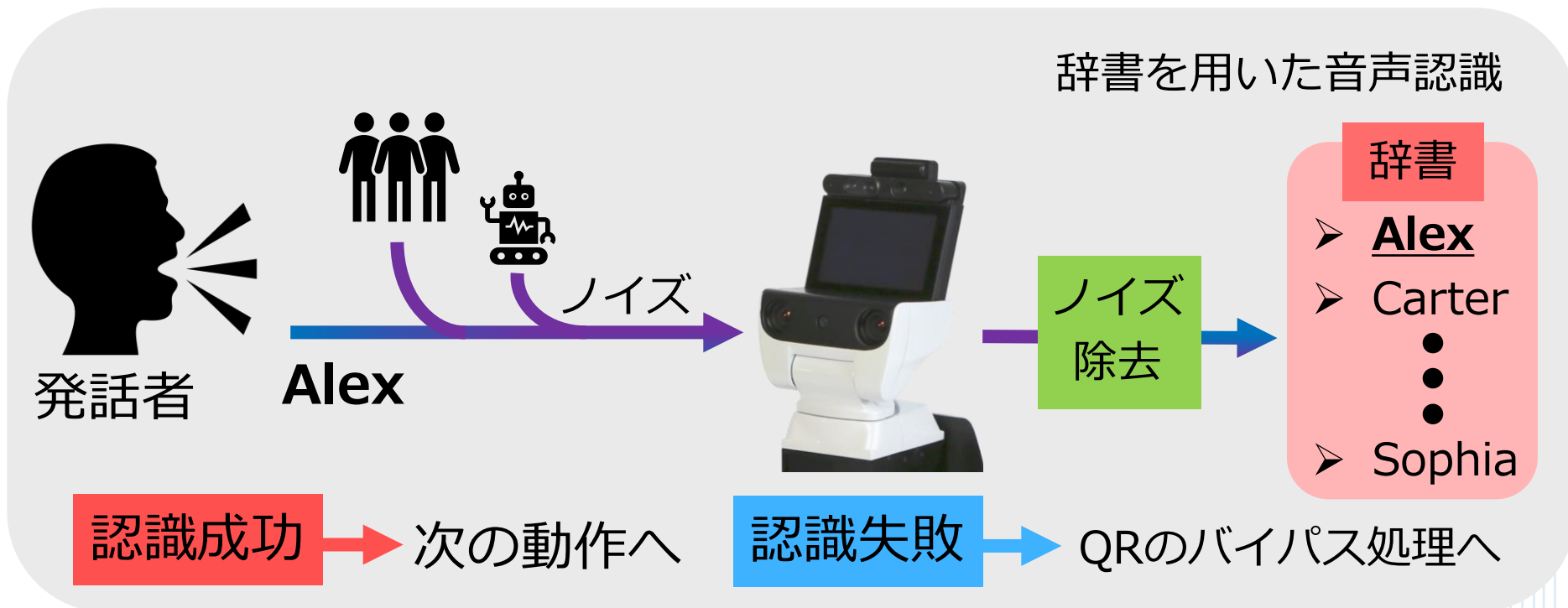
- 観客や別競技の影響で、周囲の雑音（ノイズ）が大きい
- 会場のインターネット接続が不安定である
 - ✓ ネット環境の課題は一般家庭でも発生するため、対応が必要



ノイズ除去とオフラインでの音声認識が必要

使用する音声認識手法

- 音声認識手法：Vosk^[6]
 - ✓ オフラインで動作
 - ✓ 認識する単語リスト（辞書）を設定可能
- 強度の調節が可能なノイズ除去^[7, 8]を前段に組み込む
- 認識できなかった場合, QRコードによるバイパスを使用

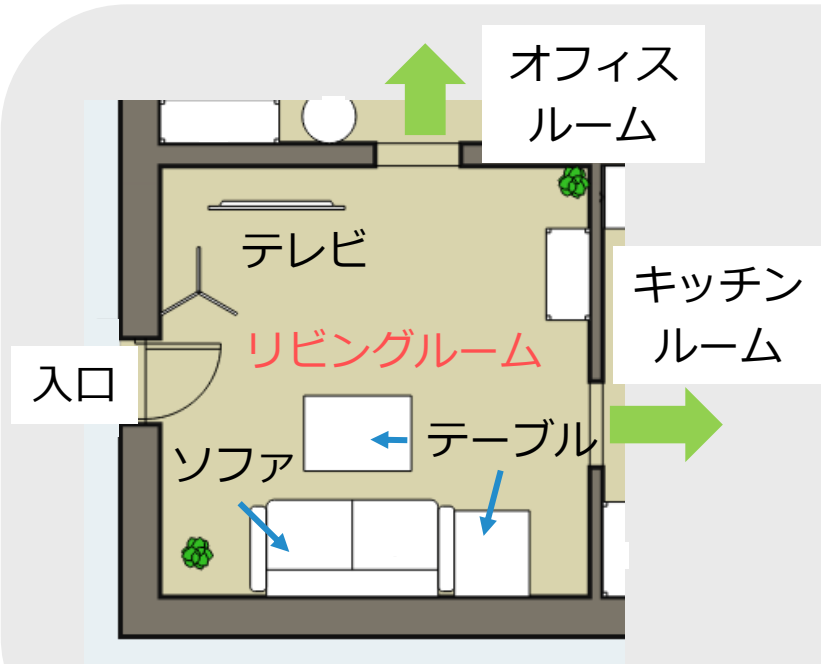


[6] vosk, <https://alphacephei.com/vosk/>, 2022 10/22 accessed

[7]T. Sainburg, et al., Plos computational biology, 2020. [8]T. Sainburg, Zendo, 2019.

競技会概要と検証内容

2022年7月にバンコクで開催されたRoboCup@Homeに参加



FMMが行われた会場の様子

検証内容

1. 会場での音声認識精度検証
2. Find My Matesを2回行い，提案手法の有効性を検証

音声認識の精度検証

検証方法

- 会場で取得した雑音のある発話データ(計88個)で検証
- 発話した文字列と認識文字列が完全一致しているかを判定

検証結果

	ノイズ除去なし	ノイズ除去あり
辞書指定なし	13.6 %	10.2 %
辞書指定あり	69.3 %	<u>71.6 %</u>

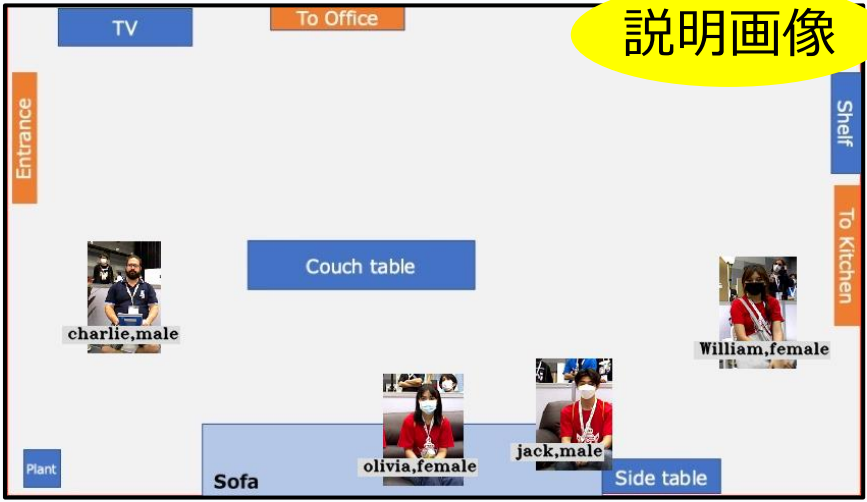
- 辞書指定・ノイズ除去を行うことで、精度が58ポイント増加
- ノイズ除去を行うことで、認識を失敗するデータもあった

競技動画（2回目）

①部屋中央までのナビゲーション



結果詳細

	1回目	2回目
1. ナビゲーション	×	○
2. 人物認識	○	○
3. 性別推定	×	○
4. 音声認識	△ (QR)	△ (QR)
5. ホストへの報告	<p>×</p> <p>ナビゲーションの不具合で 画像が不鮮明となり、 報告したと 認められなかった</p>	<p>○</p> 

2回目の試行で満点を獲得

考察：音声認識

- ノイズ除去を行ったことで、認識に失敗するデータがあった
 - ✓ ノイズ除去のパラメータを、データごとに決定する必要がある
 - ✓ 認識にとって重要な音声をノイズとして除去している可能性がある
- 競技中に音声認識ができなかった要因
 - ✓ 家具の影響でHSRと人の距離が遠く、発話内容が聞き取りにくかった



今後は、音声強調手法^[9]の導入を検討

まとめと今後の展望

RoboCup@Homeで行われるFMMへの解法を提案し
競技会を通して有効性を検証した

実験結果

● 競技結果

- ✓ 2回目の試行で、満点を獲得した
- ✓ 音声認識は結果を得ることができず、バイパスを用いることになった

● 音声認識の精度検証

- ✓ 会場でノイズのある音声データを録音して検証し、精度は71.6%であった
- ✓ ノイズ除去を行うことで、かえって認識に失敗するデータもあった

今後の展望

- 音声認識の精度向上に向けた前処理の検討
- 提案手法をより一般化し、様々なタスクに応用させる

Thank you!

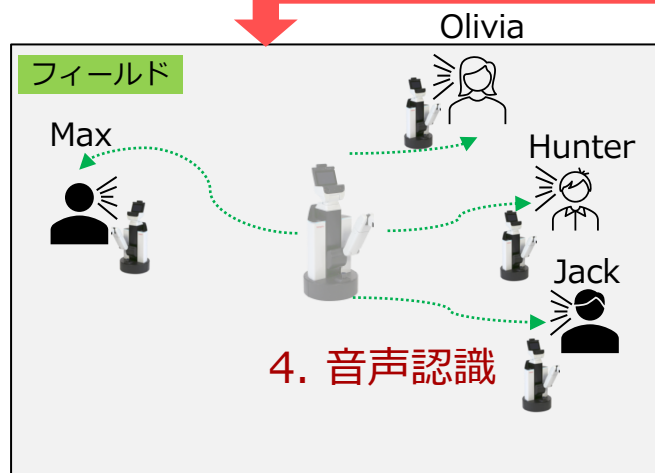
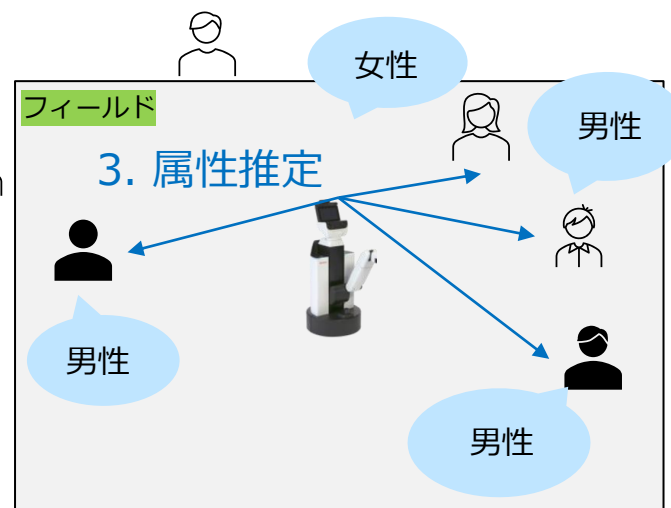
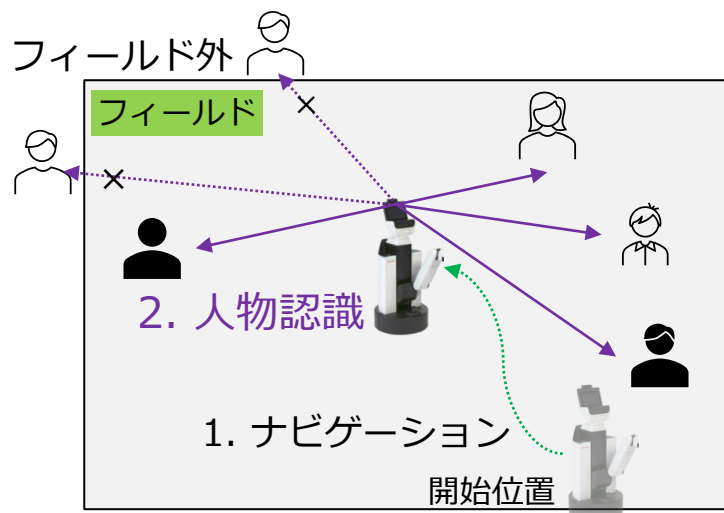
ご清聴ありがとうございました。

<http://www.brain.kyutech.ac.jp/~tamukoh/>



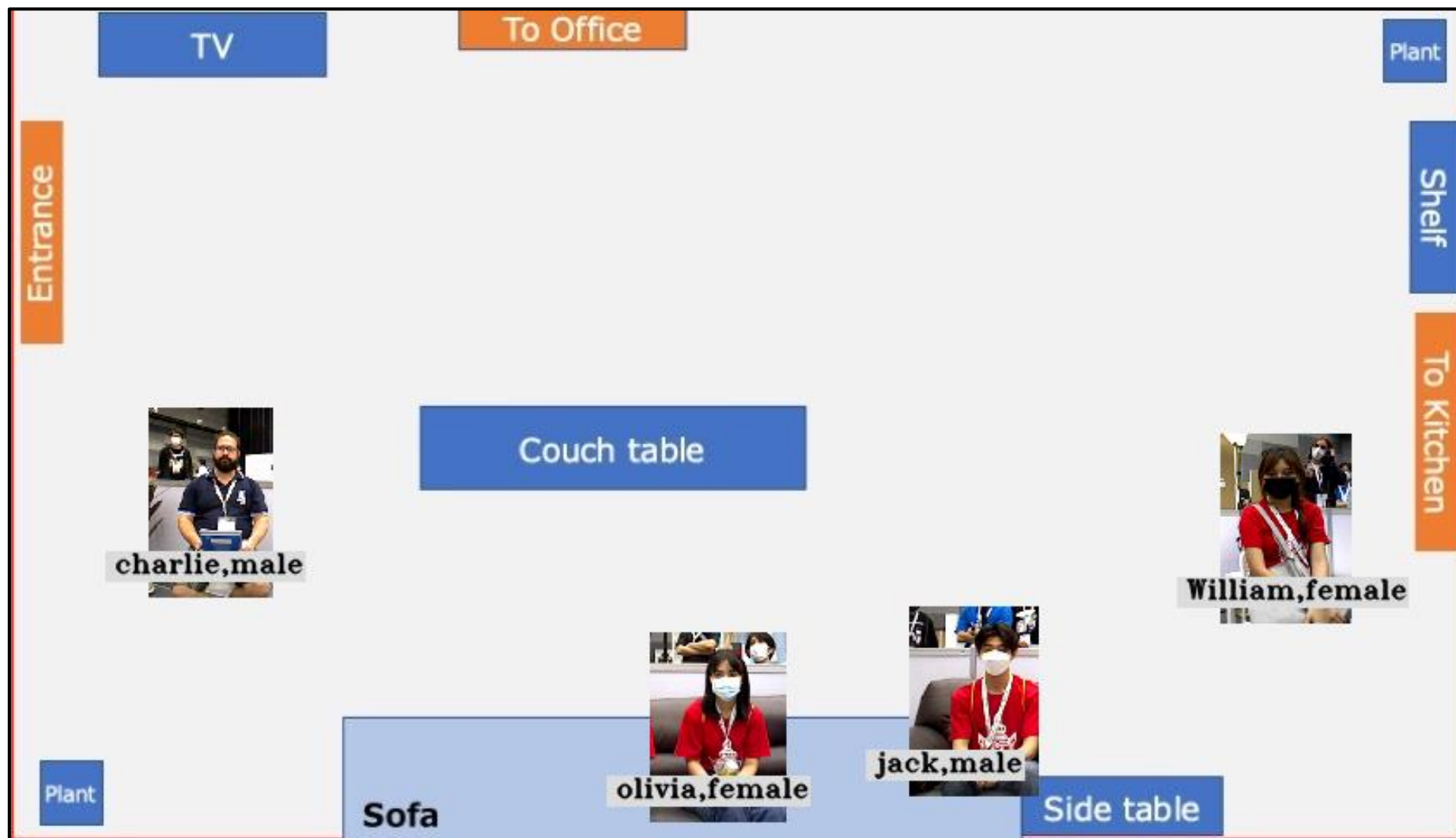
提案手法：概要

ゲストの場所・特徴を説明するための画像を作成する



説明画像により、ゲストの画像・位置・名前・性別を同時に報告する

説明画像



ゲストの画像



録音した音声のサンプルデータ

名前	ノイズ除去なし	ノイズ除去あり
Aiden		
Alex		
Brook		
Emery		
Olivia		
sophia		

実店舗でのウェイター業務

レストランで、客を見つけて注文を聞き配膳するロボット

