

Map-Free強化学習ナビゲーションの Sim2Realによる実環境での走行

○天野大輔 (明治大学), 森岡一幸 (明治大学)

背景

- 移動ロボットでは、深層強化学習によるナビゲーションが注目されている
- LiDAR や地図作成に依存しない視覚ベース自律走行手法が求められている

従来研究

概要

- 屋内環境を対象に、カメラ画像を入力とした深層強化学習モデルを構築し、自律走行に成功した
- セマンティックセグメンテーションを用いて観測表現を統一し、学習環境と実環境のギャップを縮小した

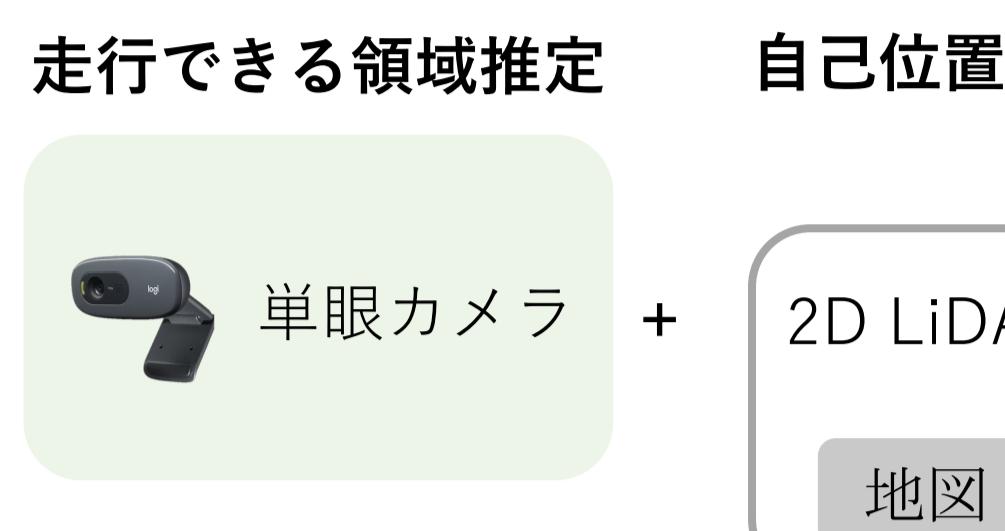
課題

- セマンティックセグメンテーションの精度が十分でなく、仮想環境との観測表現に依然ギャップが残る
- 自己位置推定が LiDAR などの高価センサに依存しており、システム全体としてカメラなどの低コストセンサのみで構成できていない
- 事前地図構築必要、未知環境への適応難しい

研究背景・目的

本研究の目的

- 事前地図なし (Map-Free) ・ LiDAR なし・カメラのみで任意の目的地へ安定走行できる RL ナビゲーションモデルを構築する
- シミュレーションで学習した方策を実機へ転移 (Sim2Real) するため観測表現を統一 (SimとRealを二値化で揃える)

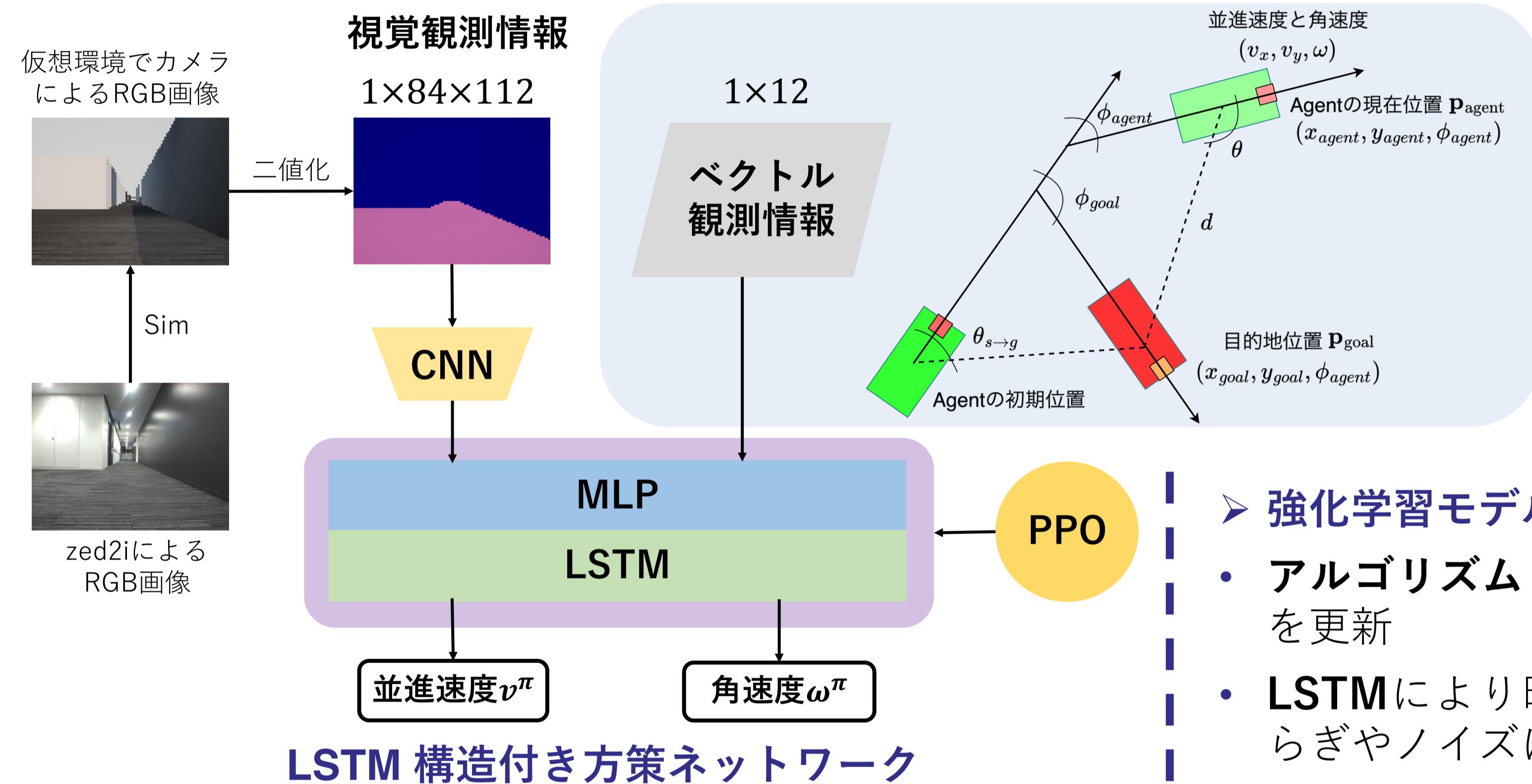


従来の手法



本研究の手法

シミュレーション環境を用いたナビゲーション方策の獲得



シミュレーション環境の二値化処理

- 床面を赤紫、壁面を暗青に着色した二値化シミュレーション環境を用いて、照明変動を受けない安定した視覚観測を実現した

観測情報

- 二値化画像と 5 種類のベクトル観測を統合して方策入力を構成した

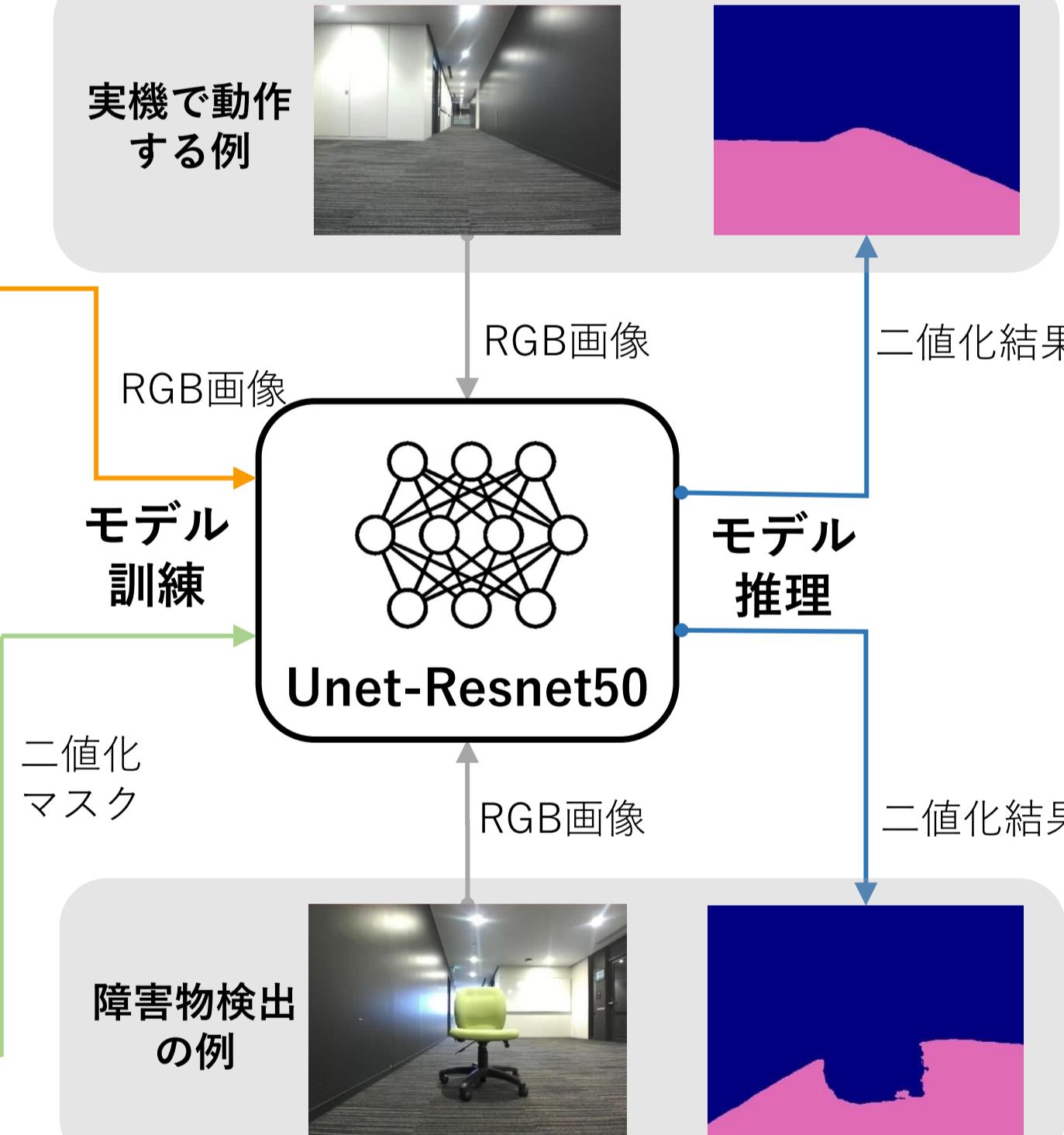
報酬設計

報酬項目	報酬値
成功報酬	+30
距離報酬	ΔJ_{nav} -20
衝突罰則	-0.01
遅延罰則 (1 ステップごとに)	-0.01
後退罰則 (負速度時)	-0.01
振動罰則 (0.2 秒以内移動 0.05 m 未満)	-0.01
切替罰則 (5 ステップ切替 3 回以上)	-0.01

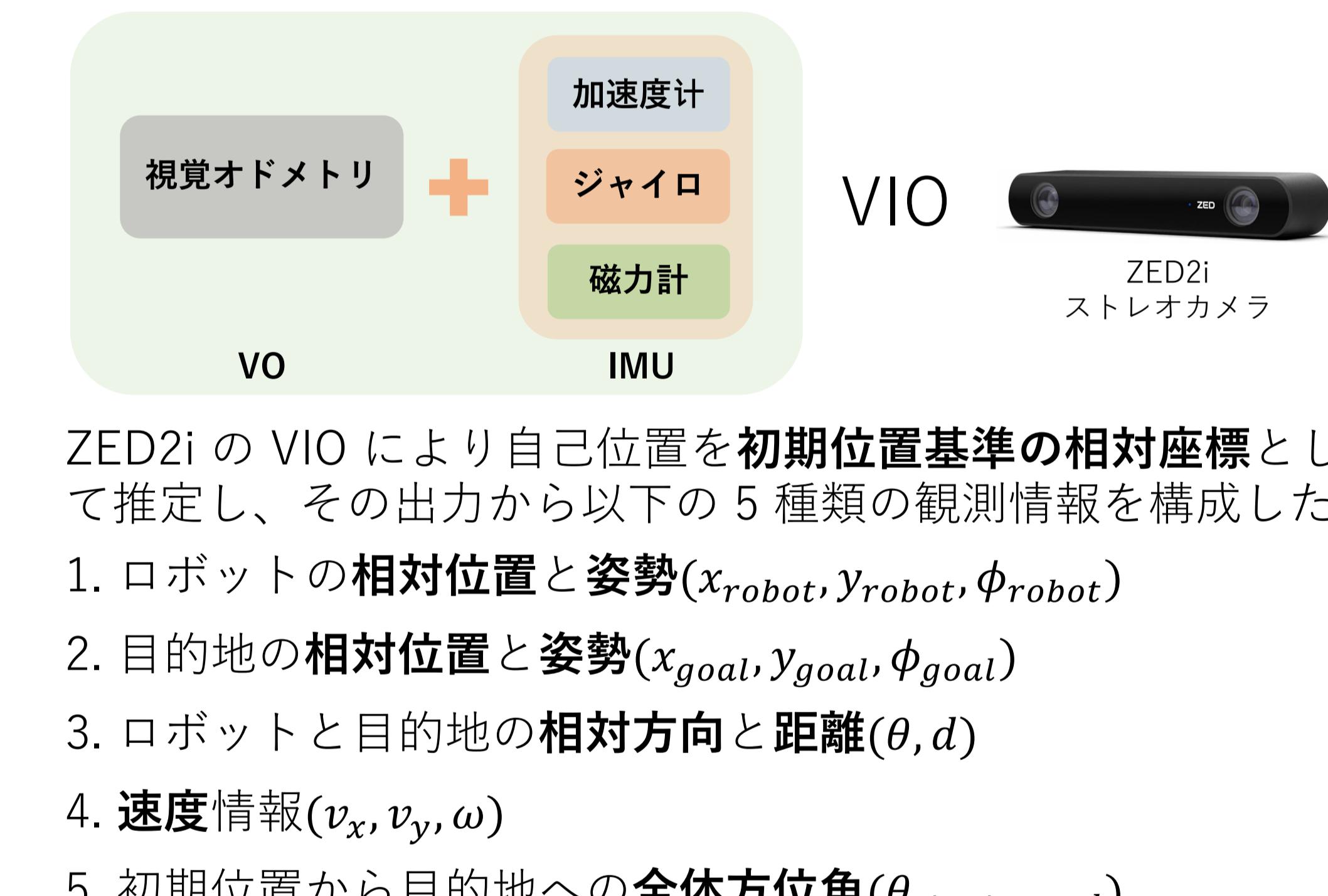
実環境への転移手法

地面二値化分割モデル

高精度モデルSAM2を用いたデータセットの構築

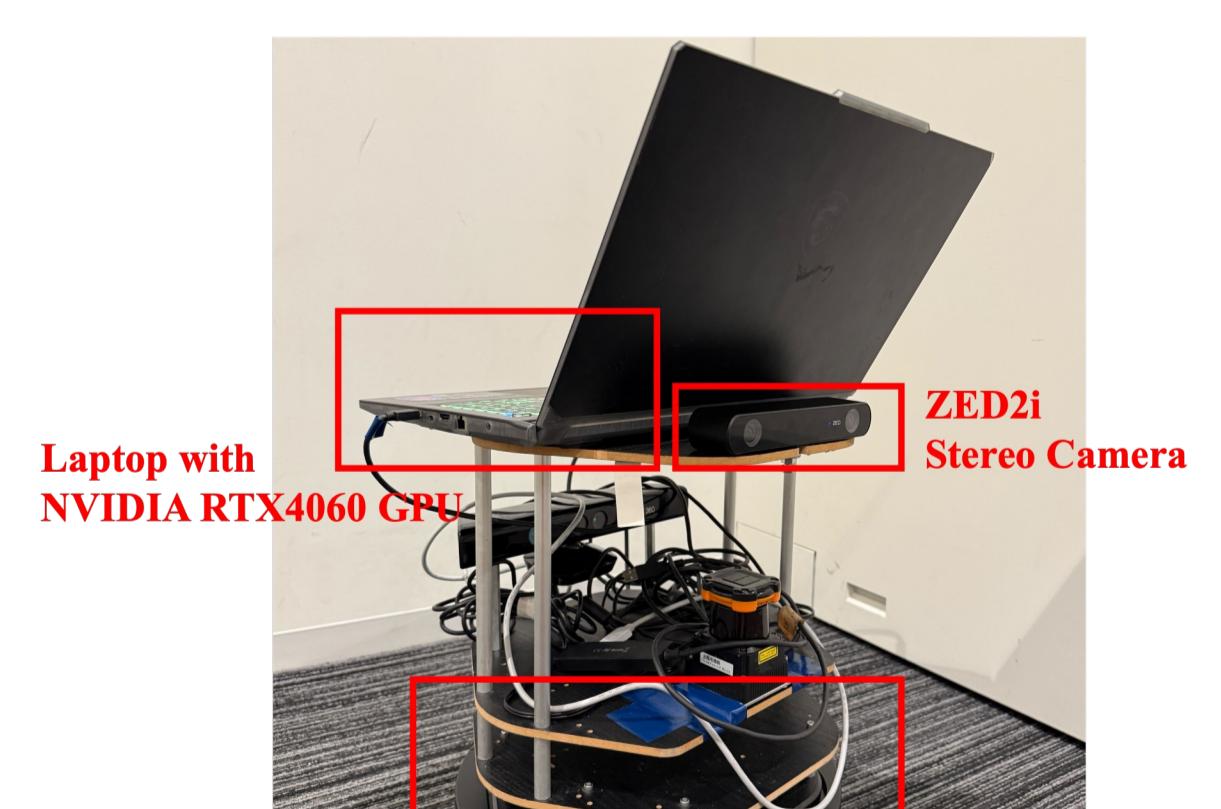


多センサー融合による自己位置推定



- ZED2i の VIO により自己位置を初期位置基準の相対座標として推定し、その出力から以下の 5 種類の観測情報を構成した：
- 1. ロボットの相対位置と姿勢 ($x_{\text{robot}}, y_{\text{robot}}, \phi_{\text{robot}}$)
- 2. 目的地の相対位置と姿勢 ($x_{\text{goal}}, y_{\text{goal}}, \phi_{\text{goal}}$)
- 3. ロボットと目的地の相対方向と距離 (θ, d)
- 4. 速度情報 (v_x, v_y, ω)
- 5. 初期位置から目的地への全体方位角 ($\theta_{\text{start} \rightarrow \text{goal}}$)

実機走行による検証実験



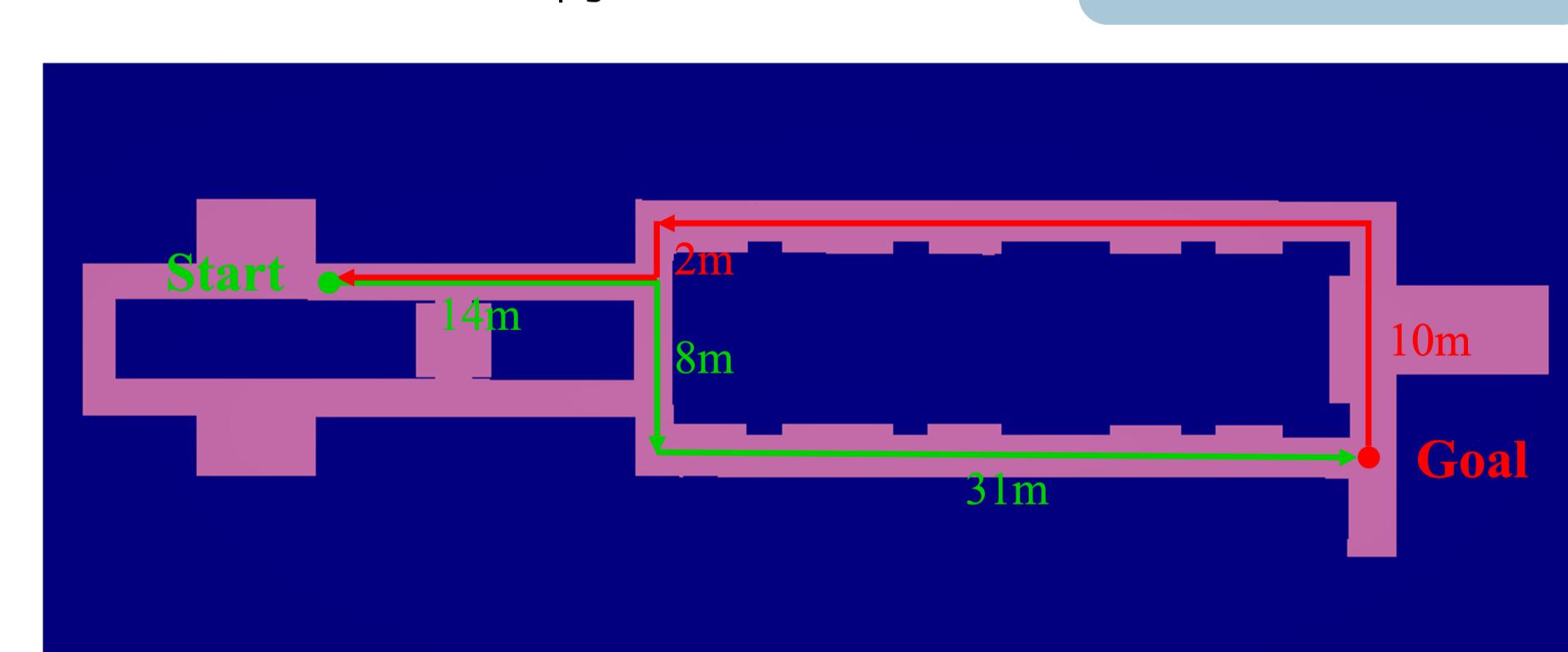
実験内容

大学廊下で約110mの往復走行

- Start → Goal : 約 53 m
- Goal → Start : 約 57 m

リモコンで一度
ゴールへ移動
ゴールの座標
を記録する

RL ポリシーに
よる自律走行



システム構成

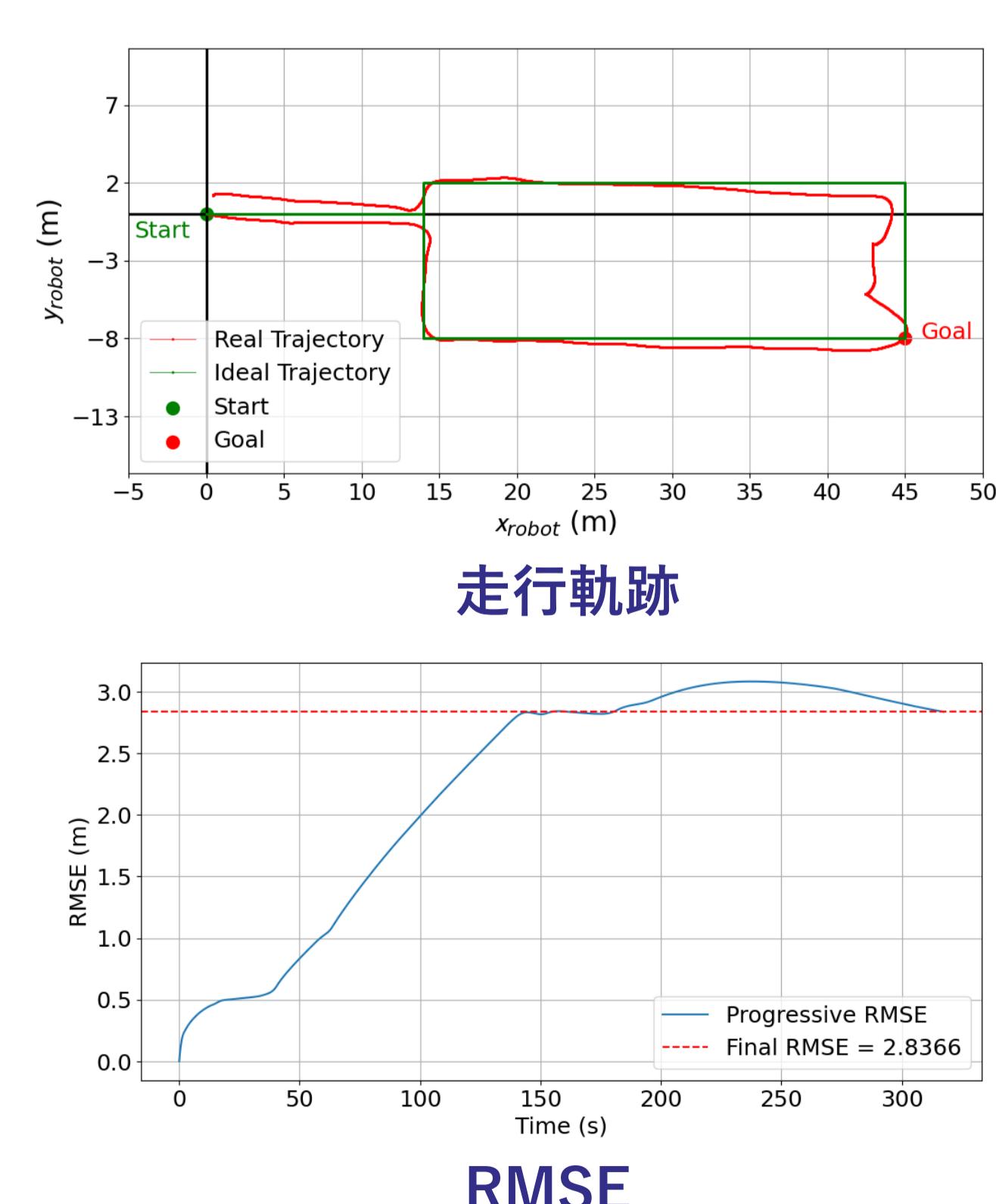
実験環境地図

実験結果

- 往復どちらも安定到達
- 曲がり角でも衝突なし
- 最終位置誤差 : 約 2 m (VIO の累積誤差が原因)
- RMSE : 約 2.84 m

課題

- VIO の累積ドリフトにより、最終位置に数メートルの誤差が生じた



まとめ・今後の展望

まとめ

- Unity 上で学習した Map-Free RL ナビゲーション方策を実機へ転移した
- SAM2 教師の地面二値化モデル+VIO により LiDAR や地図不要でも安定走行を実現した
- 長距離走行においても RL ポリシーが安定して動作することを確認した

今後の展望

- SLAM (回環検出) を導入し、長距離走行時の VIO ドリフトを低減する
- より複雑な環境や障害物回避を含む実環境ナビゲーションへ拡張する
- EKF (拡張カルマンフィルタ) によるマルチセンサ融合を導入し、自己位置推定精度をさらに向上させる