An MCTS-DRL Based Obstacle and Occlusion Avoidance Methodology in Robotic Follow-Ahead Applications

Sahar Leisiazar, Edward J. Park, Angelica Lim and Mo Chen

SIMON FRASER UNIVERSITY

Introduction

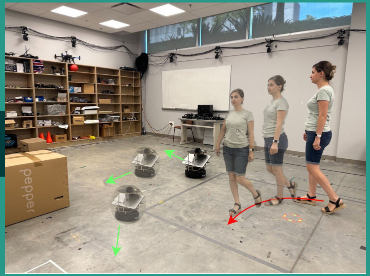
An autonomous mobile robot that follows a target person in front of them within a certain distance while ensuring avoidance of occlusion and collision avoidance.

Applications:

- Shopping carts
- Autonomous suitcases
- Recording physical activity
- Monitoring elderly people




Contribution



- *A follow-ahead mobile robot*
  - navigates in front of a target person
  - ensuring avoidance of collisions and occlusions caused by surrounding objects.

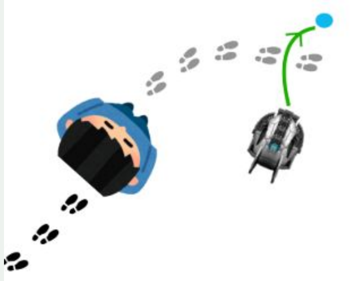
- *A high-level decision-making algorithm*
  - generates navigational goals
  - adoptes the MCTS in the context of path planning

- *Integrating Monte Carlo Tree Search (MCTS) with Deep Reinforcement Learning (DRL)*
  - enhances the performance of the decision-making
  - generates goals in any complex environments with a variety of obstacles.

Background

Previous paper (LBGP):

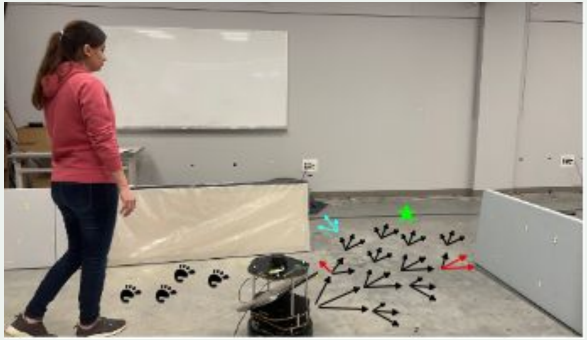
- In an obstacle free environments
- RL to estimate human trajectory and output goal point.



Nikdel, Payam, Richard Vaughan, and Mo Chen. "Lbgp: Learning based goal planning for autonomous following in front." In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3140-3146. IEEE, 2021.

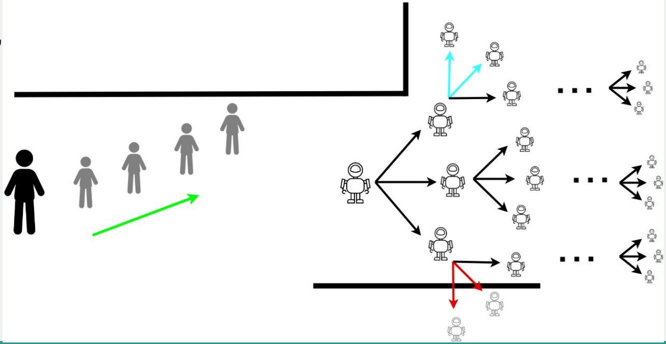
Methodology

- MCTS-RL generates goals
- TEB planner navigates the robot



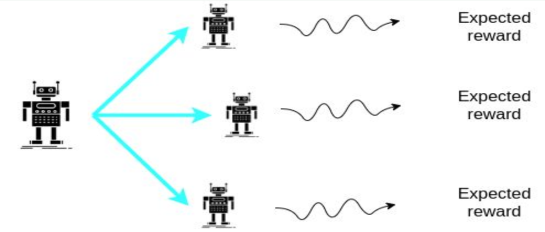
Monte Carlo Tree Search

- Inputs:
  - Human's traj prediction for 3 seconds,
  - An occupancy map of the environment,
- Expands a tree to find a best goal point for the next 3 sec,
- Considers robot and human's current and future poses as the nodes of the tree,
- Assigns the value of (-1) to a node when occlusion or collision happens,
- The value of (1) means that the robot is in front of the human,
- Selects a leaf node with the highest value as a goal point.
- Updates the goal point each $\delta t = 0.5$ sec



Reinforcement Learning

Is trained to estimate the expected return of each action and helps MCTS to evaluate each node



Return is computed as discounted sum of rewards throughout an episode

$$R = \sum_{i=0}^{n} \gamma^i r_i$$

Receives a higher reward if it stays in front of the human within certain distance.

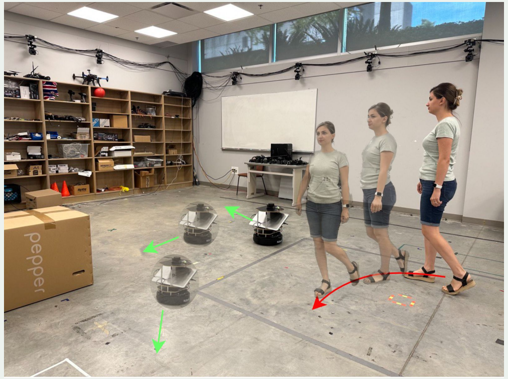
$$r = \max(r_d + r_\alpha, -1),$$

$$\text{where } r_d = \begin{cases} -(1 - d_h) & 0.5 < d_h < 1 \\ 0 & 1 < d_h < 2 \\ -0.25(d_h - 1) & d_h < 5 \quad \text{or} \quad d_h > 2 \\ -1 & \text{otherwise} \end{cases}$$

$$\text{and } r_\alpha = (45 - \alpha)/45$$

Experiments and Results

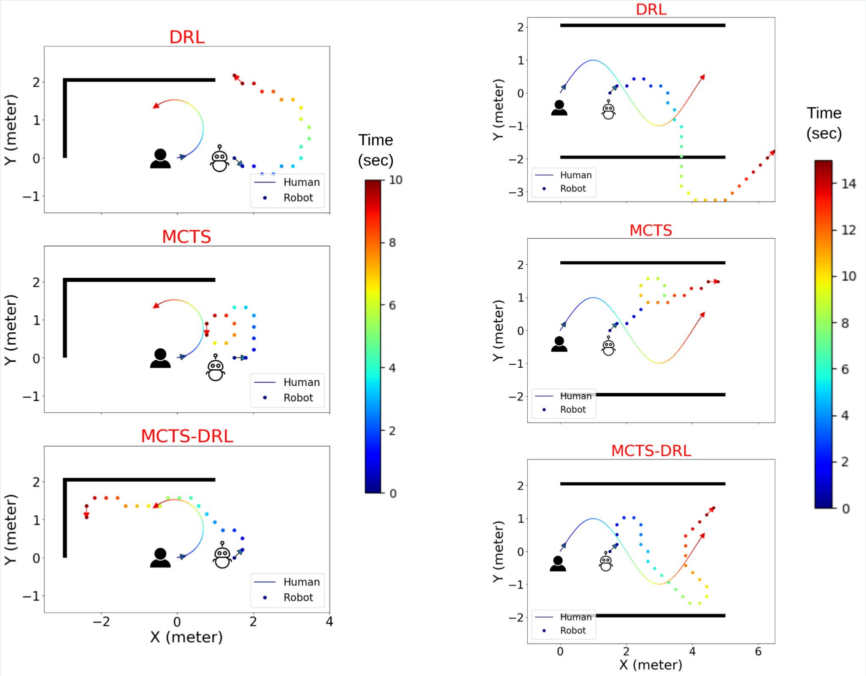
1. MCTS-DRL vs MCTS and DRL
2. Follow-ahead in an obstacle-free environment
3. Performance Evaluation with and without Obstacles



MCTS-DRL vs MCTS and DRL

- The DRL approach is unable to effectively follow the human and avoid obstacles
- The MCTS approach fails to generate consistent results, due to its random action selection
- MCTS-DRL demonstrates a superior performance
- Table shows sum of rewards through 20 trials

Human Trajectory	DRL	MCTS	MCTS-DRL
Circle	-17.95	2.87 ± 5.96	$\mathbf{4.53}$
S-shaped	-21.84	-3.83 ± 4.33	$\mathbf{-1.61}$



Follow-ahead in an obstacle-free environment

- Comparing mean distance and orientation of robot with LBGP method and ours
- No obstacle
- We achieve comparable results to LBGP

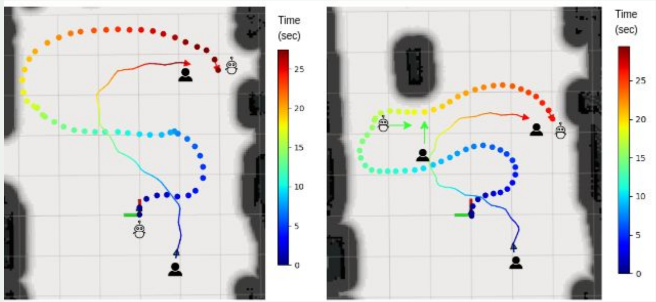
Human Trajectory	Method	Distance (m) Mean ± std	orientation(α (deg)) Mean ± std
	MCTS-DRL	**1.32 ± 0.11**	-3.47 ± 7.8
	LBGP	1.24 ± 0.3	**2.1 ± 14.5**
	MCTS-DRL	1.16 ± 0.18	**-15 ± 16.25**
	LBGP	**1.61 ± 0.4**	30.2 ± 34.6
	MCTS-DRL	1.12 ± 0.1	**4.7 ± 17**
	LBGP	**1.63 ± 0.4**	-8.4 ± 26.6
	MCTS-DRL	**1.47 ± 0.32**	**-8.76 ± 132.04**
	LBGP	1.34 ± 0.4	20.9 ± 37.8
	MCTS-DRL	**1.44 ± 0.33**	14.39 ± 119.68
	LBGP	1.99 ± 0.2	**-16.8 ± 18.6**

Human Trajectory	Method	Distance (m) Mean ± std	orientation(α (deg)) Mean ± std
	MCTS-DRL	**1.33 ± 0.34**	-28.33 ± 62.24
	LBGP	1.86 ± 0.4	**16.9 ± 28.3**
	MCTS-DRL	**1.20 ± 0.32**	-19.01 ± 49.51
	LBGP	2.09 ± 0.3	**-4.9 ± 26.7**
	MCTS-DRL	**1.21 ± 0.36**	**-28.64 ± 55.17**
	LBGP	1.81 ± 0.6	35.8 ± 33.1
	MCTS-DRL	**1.51 ± 0.2**	**12.01 ± 142.07**
	LBGP	n/a	n/a
	MCTS-DRL	**1.53 ± 0.34**	**-10.01 ± 98.97**
	LBGP	n/a	n/a

Performance Evaluation with and without Obstacles

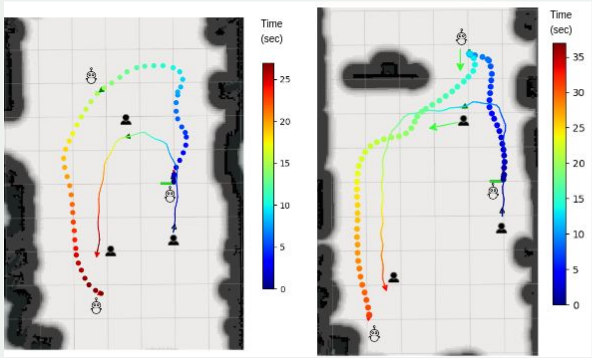
S_shape:

The robot changed its direction to avoid occlusion at T=17 sec.



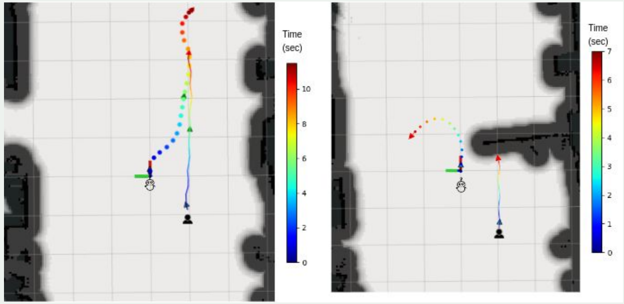
U-shape:

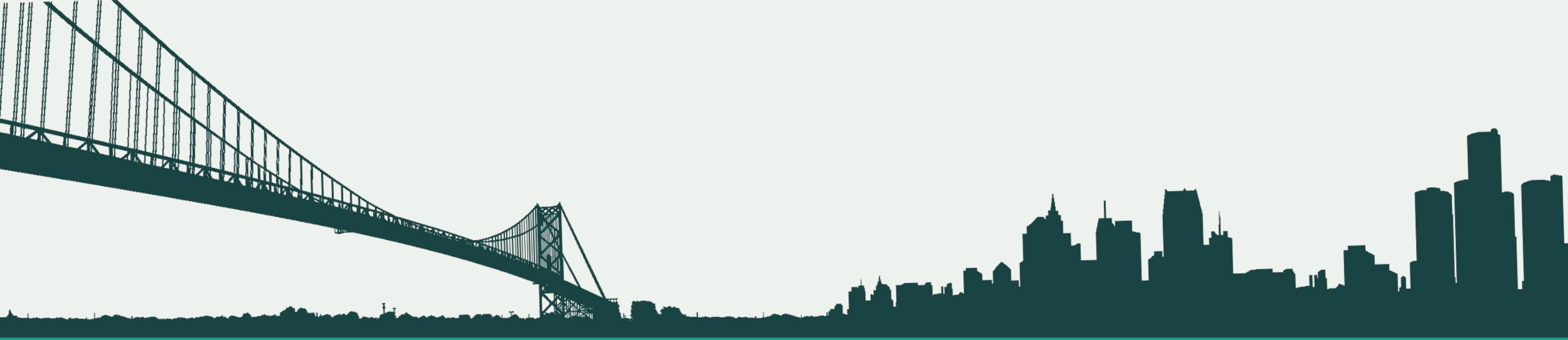
The robot changed its direction to avoid occlusion at T=12 sec.



Straight line:

The robot turns left to avoid occlusion.





Thank you

Scan to view the project description