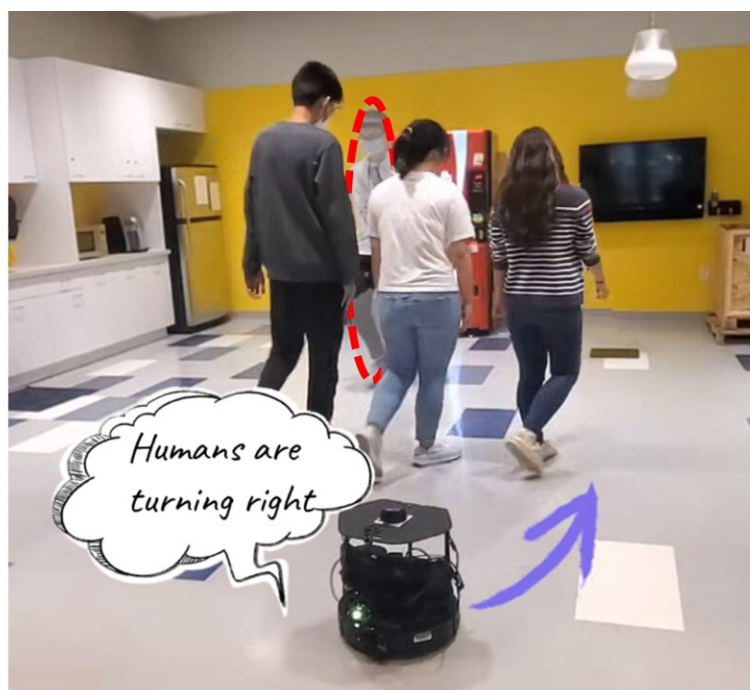


Improving the autonomous navigation of mobile robots in crowded spaces using people as sensors

November 2 2022, by Ingrid Fadelli



Turtlebot2i reasoning about a potential occluded human based on interactive behaviors of the observed humans.

Credit: Mun et al.

A team of researchers from University of Illinois at Urbana-Champaign and Stanford University led by Prof. Katie Driggs-Campbell, have recently developed a new deep reinforcement learning-based method

that could improve the ability of mobile robots to safely navigate crowded spaces. Their method, introduced in a paper pre-published on arXiv, is based on the idea of using people in the robot's surroundings as indicators of potential obstacles.

"Our paper builds on the 'people as sensors' research direction for mapping in the presence of occlusions," Masha Itkina, one of the researchers who carried out the study, told TechXplore. "The key insight is that we can make spatial inferences about the environment by observing interactive human behaviors, thus treating people as sensors. For example, if we observe a driver brake sharply, we can infer that a pedestrian may have run out on the road in front of that driver."

The idea of using people and their interactive behaviors to estimate the presence or absence of occluded obstacles was first introduced by [Afolabi et al](#) in 2018, specifically in the context of self-driving vehicles. In their previous work, Itkina and her colleagues built on this group's efforts, generalizing the 'people as sensors' idea so that it considered [multiple observed human drivers](#), instead of a single driver (as Afolabi et al.'s approach did).

To do this, they developed a "sensor" model for all the different drivers in an [autonomous vehicle](#)'s surroundings. Each of these models mapped the driver's trajectory to an occupancy grid representation of the environment ahead of the driver. Subsequently, these occupancy estimates were incorporated into the autonomous robot's map, using sensor fusion techniques.

"In our recent paper, we close the loop by considering occlusion inference within a reinforcement learning pipeline," Itkina said. "Our aim was to demonstrate that occlusion inference is beneficial to a downstream path planner, particularly when the spatial representation is task-aware. To achieve this objective, we constructed an end-to-end

architecture that simultaneously learns to infer occlusions and to output a policy that successfully and safely reaches the goal."

Most previously developed models viewing people as sensors are specifically designed to be implemented in [urban environments](#), to increase the safety of autonomous vehicles. The new model, on the other hand, was designed to improve a mobile robot's ability to navigate crowds of people.

Crowd navigation tasks are generally more difficult than urban driving tasks for autonomous systems, as human behaviors in crowds are less structured and thus more unpredictable. The researchers decided to tackle these tasks using a deep reinforcement learning model integrated with an occlusion-aware latent space learned by a variational autoencoder (VAE).

"We first represent the robot's surrounding environment in a local occupancy grid map, much like a bird's-eye view or top-down image of the obstacles around the robot," Ye-Ji Mun, the first author on this study, told TechXplore. "This occupancy grid map allows us to capture rich interactive behaviors within the grid area regardless of the number or size and shape of the objects and people."

The researchers' model includes an occlusion inference module, which was trained to extract observed social behaviors, such as slowing down or turning to avoid collisions from collected sequences of map inputs. Subsequently, it uses this information to predict where occluded objects or agents might be located and encodes this "augmented perception information" into a low dimensional latent representation, using the VAE architecture.

"As our occlusion inference module is provided with only partial observation of the surrounding human agents, we also have a supervisor

model, whose latent vector encodes the spatial location for both the observed and occluded human agents during training," Mun explained. "By matching the latent space of our occlusion module to that of the supervisor model, we augment the perceptual information by associating the observed social behaviors with the spatial locations of the occluded human agents."

The resulting occlusion-aware latent representation is ultimately fed to a deep reinforcement learning framework that encourages the robot to proactively avoid collisions while completing its mission. Itkina, Mun and their colleagues tested their model in a series of experiments, both in a simulated environment and in the real-world, using the mobile robot Turtlebot 2i.

"We successfully implemented the 'people as sensors' concept to augment the limited robot perception and perform occlusion-aware crowd navigation," Mun said. "We demonstrated that our occlusion-aware policy achieves much better navigation performance (i.e., better collision avoidance and smoother navigation paths) than the limited-view navigation and comparable to the omniscient-view navigation. To the best of our knowledge, this work is the first to use social [occlusion](#) inference for crowd navigation."

In their tests, Itkina, Mun and their colleagues also found that their model generated imperfect maps, which do not contain the exact locations of both the observed agents and estimated agents. Instead, their module learns to focus on estimating the location of nearby 'critical agents' that might be occluded and could block the robot's path towards a desired location.

"This result implies that a complete map is not necessarily a better map for navigation in a partially observable, crowded environment but rather focusing on a few potentially dangerous agents is more important," Mun

said.

The initial findings gathered by this team of researchers are highly promising, as they highlight the potential of their method for reducing a robot's collisions with obstacles in crowded environments. In the future, their model could be implemented on both existing and newly developed mobile robots designed to navigate malls, airports, offices, and other crowded environments.

"The main motivation for this work was to capture human-like intuition when navigating around humans, particularly in occluded settings," Itkina added. "We hope to delve deeper into capturing human insights to improve robot capabilities. Specifically, we are interested in how we can simultaneously make predictions for the environment and infer occlusions as the inputs to both tasks involve historical observations of human behaviors. We are also thinking about how these ideas can transfer to different settings, such as warehouse and assistive robotics."

More information: Ye-Ji Mun et al, Occlusion-Aware Crowd Navigation Using People as Sensors, *arXiv* (2022). [DOI: 10.48550/arxiv.2210.00552](https://doi.org/10.48550/arxiv.2210.00552)

Bernard Lange et al, LOPR: Latent Occupancy PRediction using Generative Models, *arXiv* (2022). [DOI: 10.48550/arxiv.2210.01249](https://doi.org/10.48550/arxiv.2210.01249)

Oladapo Afolabi et al, People as Sensors: Imputing Maps from Human Actions, *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (2019). [DOI: 10.1109/IROS.2018.8594511](https://doi.org/10.1109/IROS.2018.8594511)

Masha Itkina et al, Multi-Agent Variational Occlusion Inference Using People as Sensors, *2022 International Conference on Robotics and Automation (ICRA)* (2022). [DOI: 10.1109/ICRA46639.2022.9811774](https://doi.org/10.1109/ICRA46639.2022.9811774)

Diederik P Kingma et al, Auto-Encoding Variational Bayes, *arXiv* (2013). [DOI: 10.48550/arxiv.1312.6114](https://doi.org/10.48550/arxiv.1312.6114)

Masha Itkina et al, Dynamic Environment Prediction in Urban Scenes using Recurrent Representation Learning, *2019 IEEE Intelligent Transportation Systems Conference (ITSC)* (2019). [DOI: 10.1109/ITSC.2019.8917271](https://doi.org/10.1109/ITSC.2019.8917271)

Maneekwan Toyungyernsub et al, Double-Prong ConvLSTM for Spatiotemporal Occupancy Prediction in Dynamic Environments, *arXiv* (2020). [DOI: 10.48550/arxiv.2011.09045](https://doi.org/10.48550/arxiv.2011.09045)

Bernard Lange et al, Attention Augmented ConvLSTM for Environment Prediction, *arXiv* (2020). [DOI: 10.48550/arxiv.2010.09662](https://doi.org/10.48550/arxiv.2010.09662)

© 2022 Science X Network

Citation: Improving the autonomous navigation of mobile robots in crowded spaces using people as sensors (2022, November 2) retrieved 27 March 2023 from <https://techxplore.com/news/2022-11-autonomous-mobile-robots-crowded-spaces.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.