

Teaching Robots to Interact with Humans in a Smart Environment

Shivam Goel

School of EECS, Washington State University
shivam.goel@wsu.edu

Abstract

Robotics in healthcare has recently emerged, backed by the recent advances in the field of machine learning and robotics. Researchers are focusing on training robots for interacting with elderly adults. This research primarily focuses on engineering more efficient robots that can learn from their mistakes, thereby aiding in better human-robot interaction. In this work, we propose a method in which a robot learns to navigate itself to the individual in need. The robotic agents' learning algorithm will be capable of navigating in an unknown environment. The robot's primary objective is to locate human in a house, and upon finding the human, the goal is to interact with them while complementing their pose and gaze. We propose an end to end learning strategy, which uses a recurrent neural network architecture in combination with Q-learning to train an optimal policy. The idea can be a contribution to better human-robot interaction.

1 Introduction

As estimated by Alzheimer's association in 2017, number of individuals more than 85 years of age is projected to triple by the year 2050. Currently, 50% of older adults require assistance with daily activities of living and the annual cost of this assistance is close to \$2 trillion [Ernst and Hay, 1994]. Thus, alternate methods of healthcare are needed to preserve their quality of life. With the advancement in technology, there has been a gradual shift of focus on creating robots to provide assistance in healthcare. The fields of machine learning and psychology are in pursuit of helping elderly adults age sustainably at home [Gessl *et al.*, 2019]. In recent years, with advancement in the field of robotics and reinforcement learning there has been an active research in the direction of creating more efficient robots and subsequently leading to better human robot interaction, a research field dedicated to designing, evaluating and understanding robotics systems to be able to efficiently communicate with people. In one of the recent works by Wilson *et al.*, a multi agent robotic system was developed that assists elder adults in a smart environment [Wilson *et al.*, 2019]. In this work, the robot is in conjunction with a smart home agent. The task of the smart home agent is

to track user activities and signal the robot to offer assistance if a person needs one. However, this work currently relies on Simultaneous Localization and Mapping (SLAM) for navigation. The basic problem for SLAM is that the map of the environment should be completely known. Instead, a more robust and intelligent navigation system should be in place so that the robot can provide better assistance. In this research we propose a map less navigation method to successfully locate the human in the smart home. The goal is to generalize the navigation problem with the primary goal of locating the human. The primary construct of the idea is to leverage the information from the LIDAR sensors and visual signals (camera images) to assist in predicting the movement of the robot in the smart environment. The final goal of the robot will be to reach out to the person to offer assistance. The information from the LIDAR sensor is same everywhere, so with this technique, there will be no need to learn the map of a new environment. Thus, it will result into learning more generalized policy for elder care assistance and can be applied to every smart home environment. This research envisions creating an end to end algorithm capable of successfully navigating to human and then being able to interact with them in a complementary pose.

2 Research Problem

2.1 Problem Statement

The goal is to perform navigation in an unknown environment towards a potential human target present in the environment. We propose the use of LIDAR sensor information in the state space. In order to capture information of human pose we run a pose detection algorithm on each image obtained from the camera mounted on the robot and append the number of human poses detected in the state space. The design of reward function remains a crucial step in this setup. Reward needs to be tied to the movement of the robot as well as the number of humans detected in each frame. The goal of the robot is to find humans and if it successfully finds humans then its goal is to approach them.

2.2 Problem Formulation

Inspired by the work of [Dobrevski and Skocaj, 2018] we define the state space by a vector $s_t = [l_1 \dots l_{30}, v_t]$, where l_1 to l_{30} correspond to the LIDAR readings obtained and

v_t corresponds to the number of human poses detected in each frame. The existing approaches for mapless navigation have information about the goal, the distance to the goal, orientation of the goal or in cases of target driven visual navigation the target image is provided [Zhu *et al.*, 2017; Dobrevski and Skocaj, 2018; Lathuilière *et al.*, 2019]. In our case the problem is much more difficult, the goal's location and pose is unknown. However, goal is always a human in our case. We propose to use realtime multi person 2D pose estimation which is shown to perform pose estimation at a speed of 30 frames per second [Cao *et al.*, 2017]. The action space is defined by $a_t = [\diamond, \leftarrow, \rightarrow, \uparrow, \downarrow]$, which denote stay, turn left, turn right and go back respectively. We plan on adding a continuous action space of the pan tilt angle which will help the robot complement the gaze of the human. Reward function is the most important formulation of the Markov Decision Process (MDP). In this case the reward should be tied to the number of human pose detected since the main goal of the robot is to find the human. Additionally, the robot's local goal is to avoid obstacles on the path. Therefore, we define a reward function as $R_t = \gamma N_t + d_t + r_t$, where N_t is the number of human pose detected, d_t will be the proximity to the nearby objects (measured by the depth sensor on the robot), and r_t is defined as high negative reward for any collision and low negative reward for each action taken at a time step. The low negative reward will aid in finding the shortest possible path to reach human. In RL, model parameters are learned on sequences of state, action and rewards, termed as episodes. The goal is to learn a policy that maps the state to actions. Q-function (action-value function) is the expected future return from a state S , taking action A_t and then following a policy π . Learning a best policy corresponds to optimization problem. To solve this optimization problem, Q-function is modelled by a neural network which takes as input the state space S_t and outputs corresponding actions pertaining to the best policy. As inspired by the works of [Lathuilière *et al.*, 2019], Long short term memory networks (LSTM) [Hochreiter and Schmidhuber, 1997] will be used. LSTM being a recurrent neural network has shown to exhibit dynamic behavior for temporal sequences.

3 Research Plan

As discussed earlier, we already have a robot setup in a smart home environment [Wilson *et al.*, 2019]. However, a new learning system needs to be developed and deployed. A simulation will be programmed using openAI gym and Gazebo [Zamora *et al.*, 2016]. Algorithm will be trained on this simulator and then transfer learning will be applied to transfer knowledge to a real robot [Taylor and Stone, 2009]. If everything is successfully learned, we plan to deploy the learning system in smart robot in smart environments and perform a study with older adults to research the advocacy of robot assistance for them.

4 Conclusions and Future Work

We propose a method to perform a mapless navigation using reinforcement learning, and envision to deploy this learning

system in a real robot. This approach relies only on the sensor information and the robot learns a policy using raw features. Once the approach is successful we plan to extend the state space with audio signals, which would help the robot to search a person more efficiently and also aid the robot to gaze at the person.

Acknowledgements

The author would like to thank Dr. Matthew E. Taylor and Dr. Diane J. Cook for their guidance and support.

References

- [Cao *et al.*, 2017] Z. Cao, T. Simon, S. Wei, and Y. Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7291–7299, 2017.
- [Dobrevski and Skocaj, 2018] M. Dobrevski and D. Skocaj. Map-less goal-driven navigation based on reinforcement learning. In *23rd Computer Vision Winter Workshop*, 2018.
- [Ernst and Hay, 1994] R.L. Ernst and J.W. Hay. The us economic and social costs of alzheimer's disease revisited. *American Journal of Public Health*, 84(8):1261–1264, 1994.
- [Gessl *et al.*, 2019] AS Gessl, S Schlögl, and N Mevenkamp. On the perceptions and acceptance of artificially intelligent robotics and the psychology of the future elderly. *Behaviour & Information Technology*, pages 1–20, 2019.
- [Hochreiter and Schmidhuber, 1997] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.
- [Lathuilière *et al.*, 2019] S. Lathuilière, B. Massé, P. Mesejo, and R. Horaud. Neural network based reinforcement learning for audio-visual gaze control in human-robot interaction. *Pattern Recognition Letters*, 118:61–71, 2019.
- [Taylor and Stone, 2009] M.E. Taylor and P. Stone. Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, 10(Jul):1633–1685, 2009.
- [Wilson *et al.*, 2019] G. Wilson, C. Pereyda, N. Raghunath, G. de la Cruz, S. Goel, S. Nesaie, B. Minor, M. Schmitter-Edgecombe, M. E. Taylor, and D. J. Cook. Robot-enabled support of daily activities in smart home environments. *Cognitive Systems Research*, 54:258–272, 2019.
- [Zamora *et al.*, 2016] I. Zamora, N.G. Lopez, V.M. Vilches, and A.H. Cordero. Extending the openai gym for robotics: a toolkit for reinforcement learning using ros and gazebo. *arXiv preprint arXiv:1608.05742*, 2016.
- [Zhu *et al.*, 2017] Y. Zhu, R. Mottaghi, E. Kolve, J. J Lim, A. Gupta, Li Fei-Fei, and A. Farhadi. Target-driven visual navigation in indoor scenes using deep reinforcement learning. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 3357–3364. IEEE, 2017.