Visual Navigation in Real-World Indoor Environments Using End-to-End Deep Reinforcement Learning

Jonáš Kulhánek, Erik Derner, and Robert Babuška

May, 2021

Czech Technical University in Prague
Czech Institute of Informatics,
Robotics and Cybernetics

Visual navigation using deep reinforcement learning

problem description:

- agent (robot) moves in a 3D environment using a set of discrete actions (forward, rotate-left, ...)
- it uses an RGB camera as its only sensor
- goal is given by an image

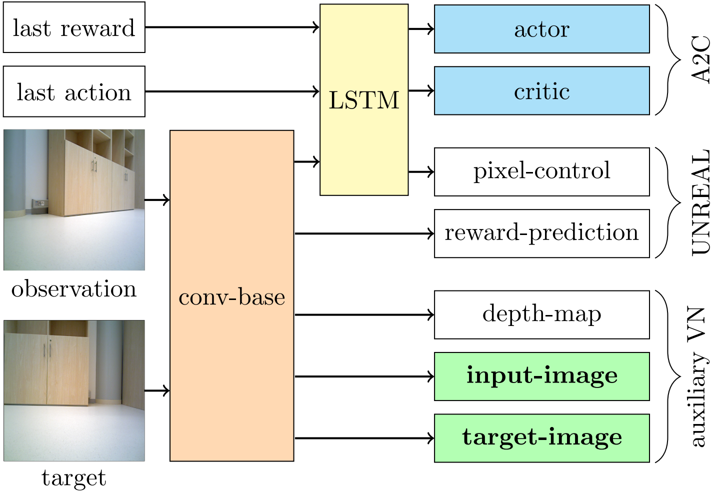
deep reinforcement learning visual navigation:

- does not need an explicit map of the environment
- current approaches applied mostly to simulated scenarios
- agent relies on the simulator to stop the navigation after the goal is reached – this cannot be used in real-world navigation

Environment simulator - DMHouse

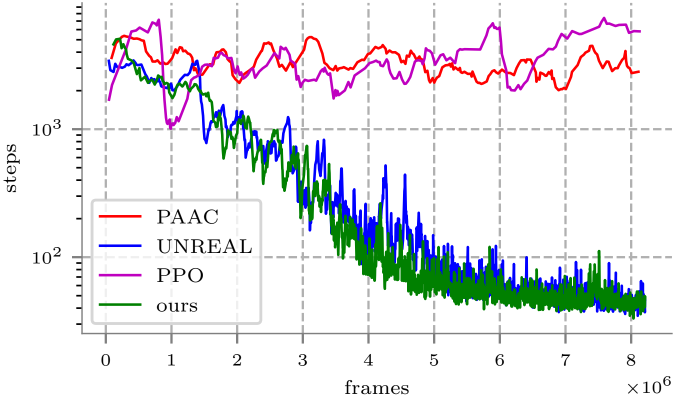
- deep reinforcement learning requires lots of images, training in real-world is not feasible
- fast 3D rendering engine designed
- capable of domain randomization (changing the placement of object, lightning conditions)

Architecture



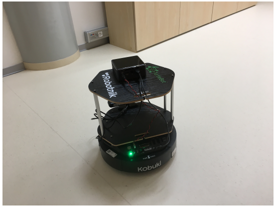
Training in simulated environment

- based on PAAC; number of parallel agents is 16
- takes approximately one day on single GPU
- PAAC and PPO fail to converge; outperforms UNREAL
- after training, 100% accuracy; avg. episode length is 42 for our method while only 46 for UNREAL



Real-world experiment

- to transfer the trained policy, we propose the following pipeline:
  1. agent is trained in DMHouse simulator
  2. images from the real-world environment are automatically collected
  3. agent is fine-tuned on the collected images
  4. agent is deployed to the real-world environment
- during training, we use curriculum learning and gradually increase the distance between the initial position and the goal
- in the real-world environment the agent (TurtleBot) reached a 0.3-meter neighbourhood of the goal in more than 86.7%



Conclusion & future work

- visual navigation using DRL, where the goal is given by an image
- novel DRL method based on PAAC enabling direct deployment of trained policy on real robots
- uses changed reward scheme and auxiliary tasks
- designed 3D environment simulator capable of domain randomization
- evaluated on real-world mobile robot, achieving satisfactory accuracy

future work:

- evaluate in larger environments with continuous motion of the robot
- perform experiments using a motion capture system for better accuracy
- improve domain randomization, allowing direct transfer