Reinforcement Learning on a Kinect-enabled Robot (SB1b-12)

Student: ZHANG, Yiyu
Supervisor: Professor Bertram E. SHI

Introduction
Reinforcement Learning (RL) has long been regarded as one of the fundamental ways to achieve the ultimate goal of robotics - being able to imitate the intelligence of human beings. In my project, I am using SARSAS, one of the popular reinforcement learning algorithms, to implement a robotic navigation system which can navigate a Kuka II robot to a desired landmark set as the destination and perform obstacle avoidance as well. A Kineci sensor is used to offer 3D depth and color information of the environment in the vision input.

Components of the System
- Inputs to the whole system
  - 3D depth image (640*480)
  - IRBA color image (640*480)
- Reward Extraction
  - Find the landmark by performing color segmentation to the color image
  - Calculate the distance between the robot and the landmark
  - Calculate the angle difference between the moving direction and the target
- Feature Values Extraction
  - Convert the depth data into a 3D point
  - Using RANSAC to perform the plane segmentation and filter out the floors and ceilings
  - Project the remaining points onto the horizontal plane and construct the occupancy grid
- Input to SARSAS
  - The state consists of the feature values and the information of the landmark
- Reward
- Output of SARSAS
  - An action to be performed by the robot
- Output of the whole system
  - A policy generating actions which can navigate the robot to the target and performs obstacle avoidance as well.

Methodology
1. System Block Diagram
   - SARSA with Linear Function Approximation
   - SARSA with Target Value Function Approximation

2. Implementation Details
   - Plane Segmentation and Occupancy Grid
   - RANSAC is used to find the plane coefficients for the inlier points
   - Floors and ceilings can be filtered out
   - Divide the polar plane into 10*50 cells with size 5.7 degree * 52 cm
   - Project the points onto the cells and count the number of points in each cell
   - Construct the occupancy grid
   - Landmark Finder
     - Find the location of the landmark using color segmentation
     - Calculate the distance and angle
   - Feature Vector
     - Feature vector into a vector
   - Build Feature Vector
   - SARSA with Linear Function Approximation
   - Action Space
     - Forward, Backward, Stop, Left, Right
   - State Space
     - Feature vector (502 elements)
     - Parameter Vectors
     - 5 vectors (302 elements) for 5 actions

Results
1. Landmark Finder
   - The left image is the result after color filtering and the right image is the locating result.

2. Plane Segmentation for Floor Filtering
   - The left image is the result of plane segmentation by using RANSAC. The right image is the result by using a straight forward way based on histograms of estimated indoor normal vectors.

3. SARSA Simulation Results (empty occupancy grid)
   - The left graph is the reward of each step and the right graph is the angle difference in each step.

4. Conclusion
   - The simulation results prove the correctness of the algorithm and the policy it generated.
   - The real experiments on robots has some problem due to the limitation of the landmark finder.