Modular Inverse Reinforcement Learning on Human Motion

Shun Zhang, Matthew Tong, Mary Hayhoe, Dana Ballard
Department of Computer Science, Center for Perceptual Systems
University of Texas at Austin

Introduction

blahh

Multi-objective Domain

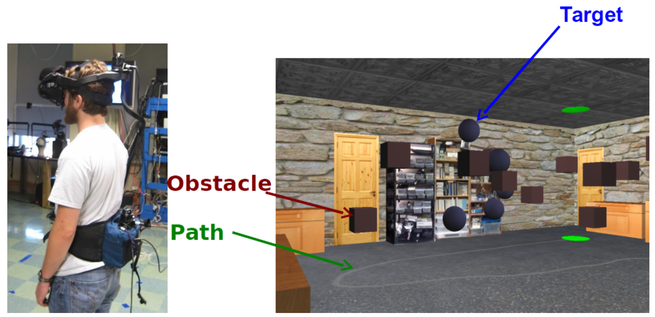
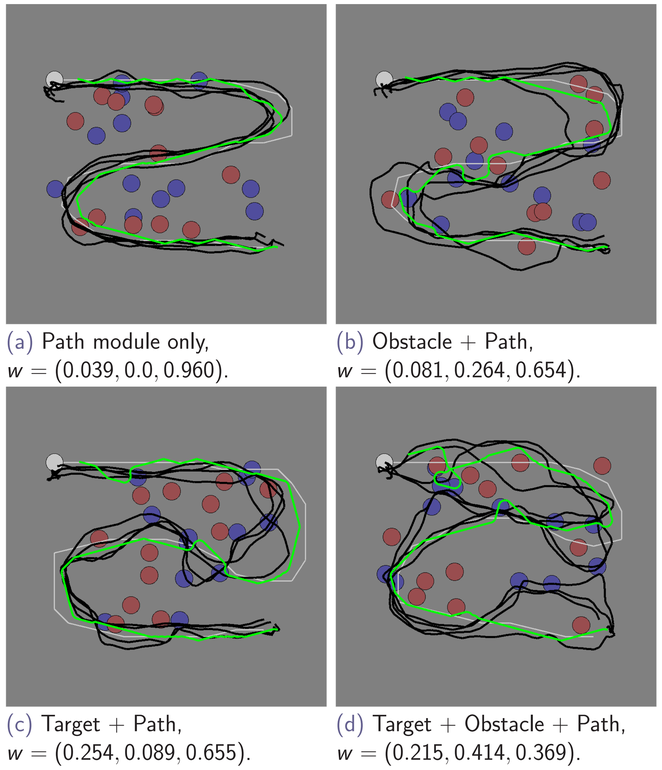


Figure : (Left) A human subject with a head mounted display (HMD) and trackers for the eye, head, and body. (Right) The environment the human can see through the HMD. The red cubes represent obstacles. The blue balls represent targets. There is also a gray path on the ground that the human subject can follow.

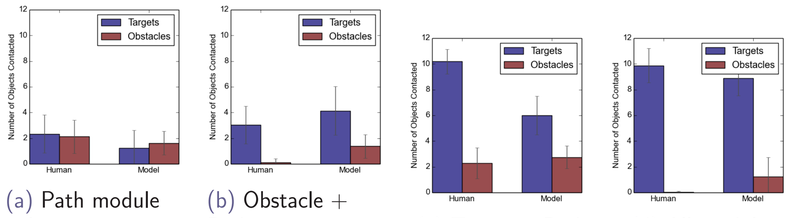
Experiments



(a) Path module only, $w = (0.039, 0.0, 0.960)$.
(b) Obstacle + Path, $w = (0.081, 0.264, 0.654)$.
(c) Target + Path, $w = (0.254, 0.089, 0.655)$.
(d) Target + Obstacle + Path, $w = (0.215, 0.414, 0.369)$.

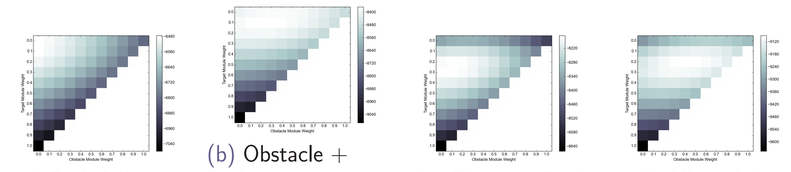
Figure : The trajectories of humans and the agent in the four tasks. Targets are blue and obstacles are red. The black lines are trajectories of human subjects, and the green lines are trajectories of the learning agent by using the optimum weights, w, derived from modular inverse reinforcement learning. Weights for each task are given as (target, obstacle, path).

Experiments



(a) Path module only.
(b) Obstacle + Path.
(c) Target + Path.
(d) All modules.

Figure : Number of targets hit and number of obstacles hit of the human subjects and the agent.



(a) Path following.
(b) Obstacle + Path.
(c) Target + Path.
(d) All modules.

Figure : Heatmaps of the log of the values of Equation **??** for different weights for the four tasks, respectively. The white zones indicate higher probabilities. The weights of all three modules sum to 1, so we only show the weights on the target and the obstacle modules.