Learning Dexterous In-Hand Manipulation

OpenAI
Marcin Andrychowicz, Bowen Baker, Maciek Chociej, Rafał Józefowicz, Bob McGrew, Jakub Pachocki, Arthur Petron, Matthias Plappert, Glenn Powell, Alex Ray, Jonas Schneider, Szymon Sidor, Josh Tobin, Peter Welinder, Lilian Weng, Wojciech Zaremba

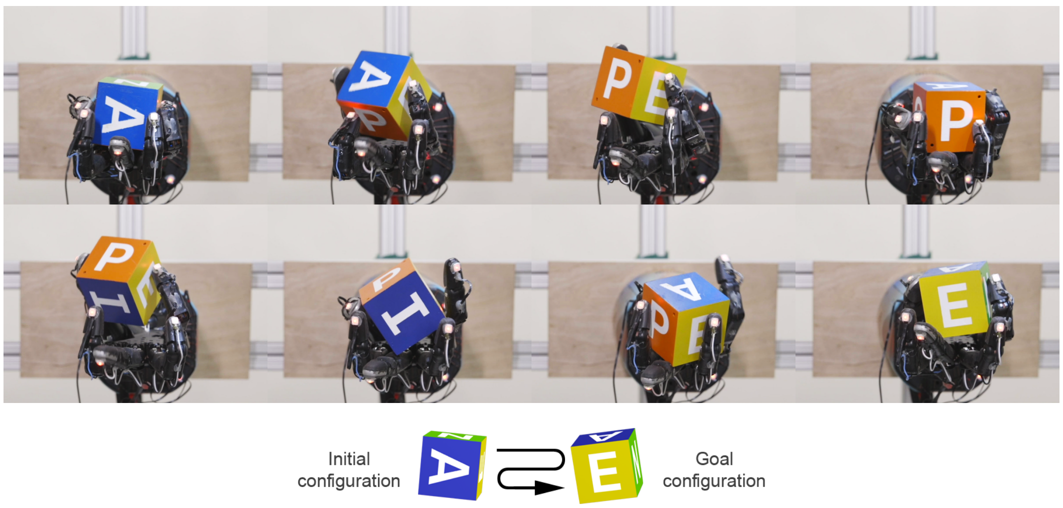

Figure 1: A five-fingered humanoid hand trained with reinforcement learning manipulating a block from an initial configuration to a goal configuration using vision for sensing.

Abstract

We use reinforcement learning (RL) to learn dexterous in-hand manipulation policies which can perform vision-based object reorientation on a physical Shadow Dexterous Hand. The training is performed in a simulated environment in which we randomize many of the physical properties of the system like friction coefficients and an object's appearance. To deal with partial observability of the environment we use recurrent policies with memory (LSTM). Our policies transfer to the physical robot despite being trained entirely in simulation. Our method does not rely on any human demonstrations, but many behaviors found in human manipulation emerge naturally, including finger gaiting, multi-finger coordination, and the controlled use of gravity.

1 Video abstract

We encourage the reader to watch a short video describing our work: `https://youtu.be/jwSbzNHGflM`.

2 System architecture

Preprint. Work in progress.

Figure 2: System Overview. (a) We use a large distribution of simulations with randomized parameters and appearances to collect data for both the control policy and vision-based pose estimator. (b) The control policy receives observed robot states and rewards from the distributed simulations and learns to map observations to actions using a recurrent neural network and reinforcement learning. (c) The vision based pose estimator renders scenes collected from the distributed simulations and learns to predict the pose of the object from images using a convolutional neural network (CNN), trained separately from the control policy. (d) To transfer to the real world, we predict the object pose from 3 real camera feeds with the CNN, measure the robot fingertip locations using a 3D motion capture system, and give both of these to the control policy to produce an action for the robot.