Computer Science and Automation

Stochastic Robotics Lab

S R L

Narendra Summer Internship 2022
Poster Presentation

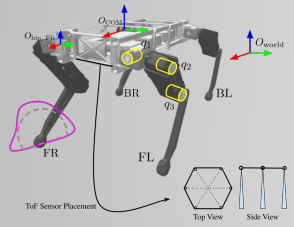
Microsoft Research

Indian Institute of Science
भारतीय विज्ञान संस्थान

Sample Efficient Model Based Reinforcement Learning for quadruped locomotion

Somnath Sendhil Kumar, Aditya Shriwatkar, Bharadwaj Amrutur, Shalabh Bhatnagar, Shishir Kolathaya

Quadruped Locomotion

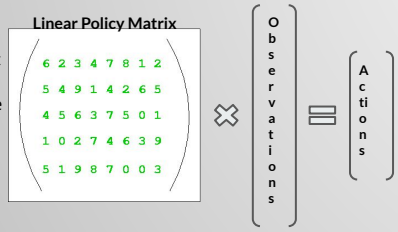


- Powered by servo motors
- Uses limited sensing like IMU, joint angles etc
- Compact design

The following is been simulated in Nvidia Isaac Gym for experimentation and testing of policies before deployment.

Augmented Random Search

Due how the problem is broken down into linear set of actions we take advantage of linearizing the actions with respect to observations, This reduces the search space and increases the convergence rate.

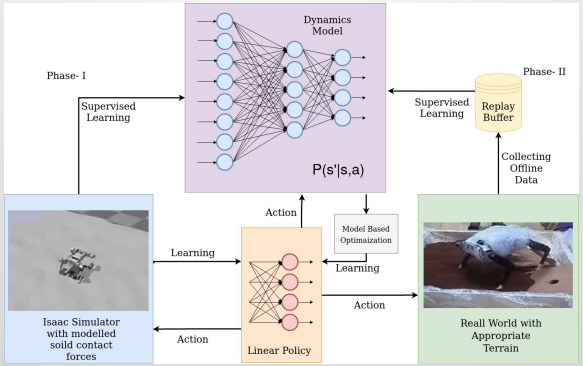


Augmented Random Search is a Evolutionary Strategy which generates a pool of agents from which the most fit candidate is selected based on reward given by the MDP Markov Decision Process.

Real world Training

Due to the Sample Inefficient nature of Evolutionary Strategies, It is very difficult to train policies directly from hardware data. There have been extensive studies to increase the efficiency of Reinforcement Learning algorithms and very few work focus on Training on Robots directly from hardware.

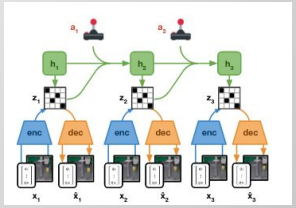
Model based training are very popular due to the fact that we can use the model for multiple Rollouts and the hardware remains safe from all the trial runs. Also Models can be used for interpreting our action and observation space and their conditional trajectories.



Results

The current Pipeline has successfully been deployed on the hardware and trained for simple walking on Flat, Slope terrain and Loose Terrain. The algorithm here has proven to be beneficial as loose terrains are impossible to simulate in any simulator due to the high number of particles and complexity in modelling the toe dynamics.

Future Work



Source : Hafner et. al.

Our current work uses model for adapting the new dynamics of the system, but we could also use a Model for learning a policy from scratch which has been done for different problem statements in the Literature.