COMP 596 - Brain-Inspired Artificial Intelligence

LE, Nhat Hung
ID: 260793376

Prof. Blake Richards
Winter 2020

Continuous Locomotion on Irregular Terrain

Humans are good at walking. Specifically, we are able to quickly adapt to rough natural terrains and navigate without losing balance. For example, a regular person can traverse a rocky trail with relative ease, and, with enough practice, even run on steeper and narrower terrains such as mountain ridges. The most advanced AI robots can now walk on smooth outdoor surfaces e.g. dirt, snow, as demonstrated by Boston Dynamics' bipedal robots, but would certainly stumble on rough, irregular terrains like the aforementioned rocky trail. This begs the question, how can real-time locomotion over irregular terrain be achieved for a (humanoid) AI? This will be analyzed at the computational, algorithmic, and implementation level.

At the computational level, locomotion over irregular terrain must be optimized for both the agent and the external environment being traversed. In other words, while walking, the agent must make a trade-off between the efficiency of the walk i.e. forward speed, and foot placement on viable locations to support continuous locomotion i.e. balance. At each step, the walker must compute (1) where to place their foot as to support their body weight, (2) how to adjust their form to maintain balance e.g. raise arms, lean body to the side, (3) how much force and in which direction to apply on the placed foot to redirect their momentum to the next viable foothold, and (4) perform a two-to-three-step lookahead in the environment to guide future foot placements.

At the algorithmic level, the first three computations can be learned through reinforcement learning. In real-time, the agent processes observations of the environment (these observations function as the aforementioned fourth computation), based upon which it then performs an action, and finally receives either a positive or negative reward based on said action. Intuitively, this is how children learn to walk early in life. Consider a policy network that inputs observations and outputs an action for each step. As previously mentioned, the agent must optimize for both speed and balance. Therefore, the policy network can consist of two subnetworks, one receiving internal or egocentric observations e.g. joint angles and velocity, the other receiving external observations of the terrain e.g. current position of agent and a representation of a limited range of the terrain ahead. As for rewards, forward progress will be rewarded, and excessive rotations, torques and tilts will be penalized.

Finally, training can be simulated in a 3D environment. The humanoid agent will possess a 3D body proper joints and weight distribution. To expose our agent to a diverse set of irregular terrains, the latter will be procedurally generated. As the agent interacts with the environment, two sets of observations are generated: (1) information local to the body such as joint angles, velocity, acceleration, physical contact on feet, etc. and (2) external information on the environment such as the position of the agent on the terrain and a limited representation of the terrain ahead. Each of the two sets of observations will be processed by its respective policy subnetwork.

In conclusion, a humanoid AI can be taught to walk on irregular terrain through reinforcement learning. In a simulated 3D environment of procedurally generated irregular terrains, the agent learns viable foot placements and forms based on two sets of observations: information local to the body, and information related to the environment. Therefore, the agent optimizes for both basic locomotor skills, and terrain perception and navigation.