Deep Visual Foresight for Planning Robot Motion, Finn et al

Summary

Massive amounts of human supervision, in terms of labeling data, is required for a robot to learn over a wide range of tasks and environments. It is imperative that a robot be capable of generating its own training data and learn from it without requiring supervision, to have a learning policy that is robust and easily generalizable. Moreover, the authors argue that a fully model-based approach, involves developing a very detailed and complex model, and can be prohibitive when dealing with a wide array of tasks and environments. The authors propose a novel method using a model-based RL, that addresses the problem described above.

Approach and Strength

The algorithm uses two models together, the deep video prediction model and a model-predictive control, to decide on what actions to perform to achieve the desired state. The models are individually described below -
Input – Current Image from the camera sensor and current pose of the robotic arm.

User Input – The user specifies a pixel to the agent as a start position, and also its desired end position. To clarify, the the user selects a pixel that overlaps an object that needs to be moved, and the robotic arm then *pushes* the object, to the requested end state.

Deep Video Prediction Model – The model uses a convolution LSTM to predict an approximate mean of the distribution of pixels of the next frame conditioned on the current image, state and future actions of the robotic arm.

Model-Predictive Control -
Due to the *deep video prediction model* predicting the probability distribution of the pixel given a set of actions, MPC uses the these pixel flow predictions for *X* number of time-steps to predict the best state of action-sequences that results in the optimal position of the pixel towards its desired end goal state.
After every time-step, the current position of the pixel is updated due to the undertaken action, and MPC is iterated again, until the pixel reaches the desired state.
A major advantage of this approach is is that no explicit definition of an object is required here, which allows the algorithm, to generalize across a large set of objects for pushing applications.

Weakness

The algorithm is tested on a simplified set of tasks, although, the authors acknowledge that in the paper. Moreover, their baseline do not use a fully-modeled algorithm that includes physics or object definitions.

Future Work and Applications

Predictive models as used in this paper, is still quite nascent. However, the results from this approach is highly interesting and can be extended to applications such as robotic navigation, trajectory planning etc.

Image prediction using GANs and other state-of-the-art techniques in image prediction can be incorporated in this approach to improve results.

Note

The authors claim that a model-free method are more suited for a task-specific policy and suffer with a lower performance when it is generalized. However, A. Tamar et al, in the paper *Value Iteration Networks*, suggest that errors in modeling a model-based learning policy can degrade the performance and hence a model-free approach is preferred.