From pixels to physics: an image-computable model of physical predictions

Haoliang Wang
MIT, Cambridge, Massachusetts, United States

Nishad Gothoskar
MIT, Cambridge, Massachusetts, United States

Joshua Tenenbaum
MIT, Cambridge, Massachusetts, United States

Kevin Smith
Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

Abstract

Having reasonable expectations of how scenes will unfold is crucial in our life. One prominent hypothesis is that we perform probabilistic physics simulation in our mind. However, the question of how people infer underlying scenes from observations and how this affects downstream predictions about physics interactions is under explored. Current models usually make simplified assumptions that the 3D geometric states of objects are already given. To better understand the role of perceptual uncertainty in people's physical predictions, we explore the idea of vision as inverse graphics and design a model that can infer a posterior distribution of object states given the raw visual inputs. This perceptual uncertainty is then propagated to a probabilistic physics simulation model to derive physical predictions. We compared the model's predictions and generalizations on a wide range of physical scenarios from the Physion dataset and found that it captured both participants' successes and error patterns.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*