Galileo: Perceiving Physical Object Properties by Integrating a Physics Engine with Deep Learning

Jiajun Wu*[1], Ilker Yildirim*[12], Joseph J. Lim[1], William T. Freeman[13], Joshua B. Tenenbaum[1]

1 Massachusetts Institute of Technology 2 The Rockefeller University 3 Google Research (* indicates equal contributions)

Motivation

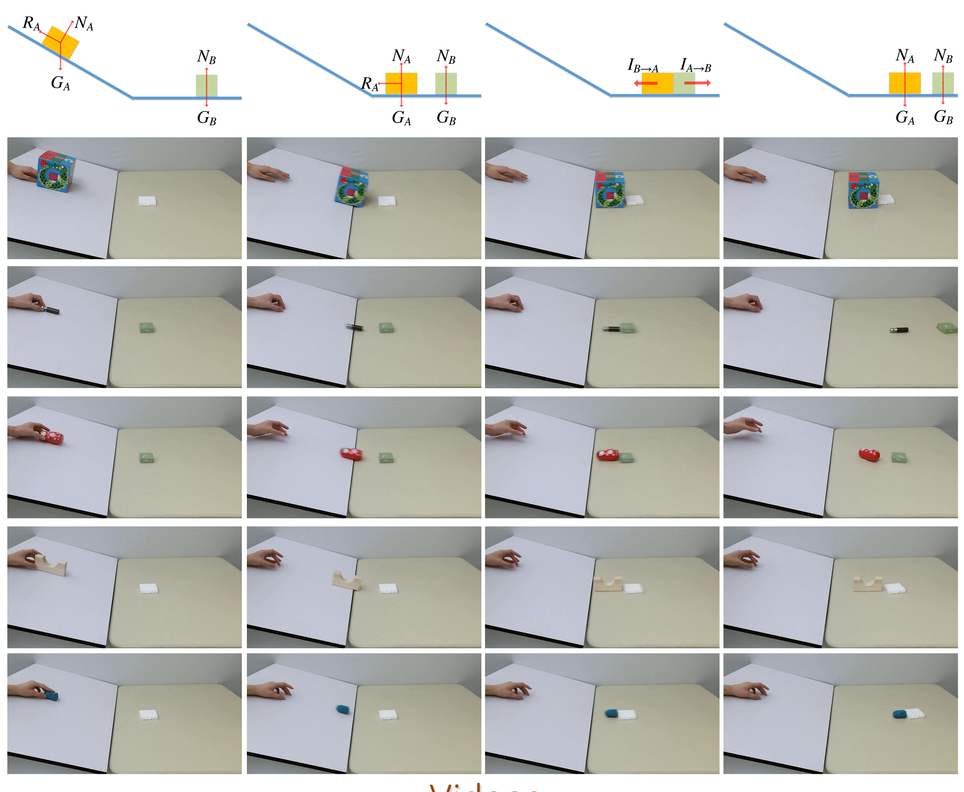
Observation

- Human infer rich information from simple visual inputs.
- Infants learn intuition on physics when they are young.
- Humans utilize a realistic physics engine as part of a generative model to interpret real-world physical scenes.

Goal: building a generative model to interpret physical scenes

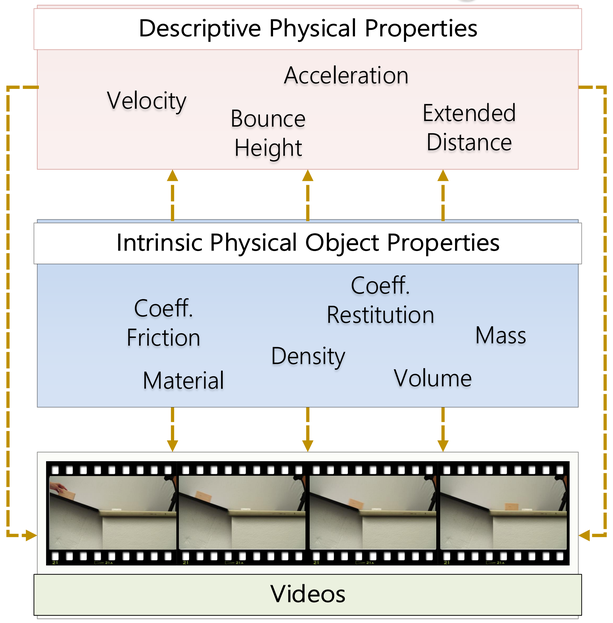
A Video Dataset

101 Objects from 15 categories

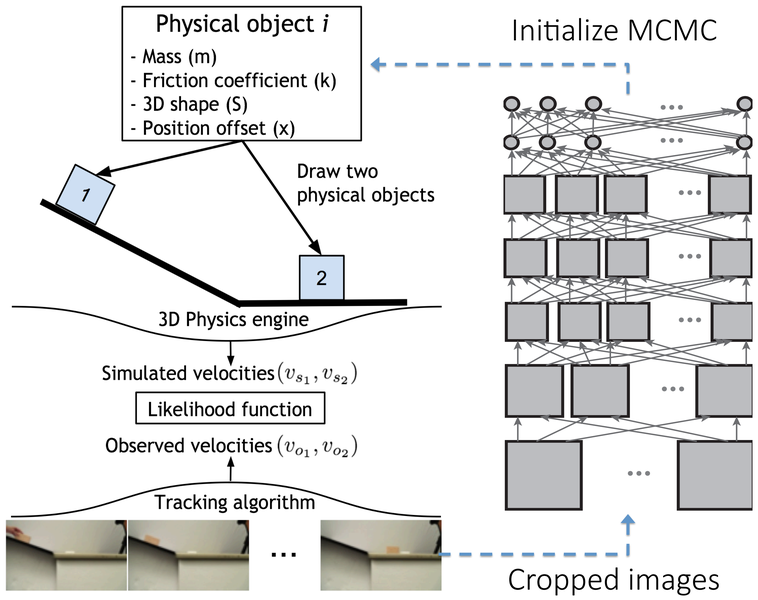


Videos

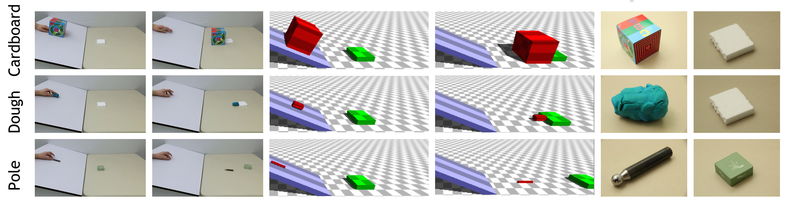
Scene Modeling



The Galileo Model

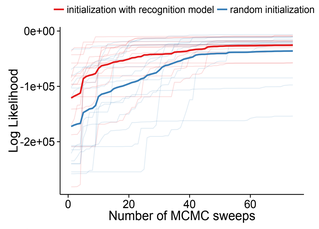


Experiments



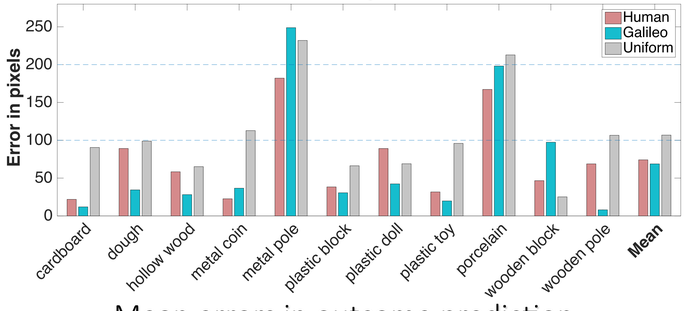
Simulation results

Method	MSE	Corr
Oracle	0.042	0.77
Galileo	0.052	0.44
Uniform	0.081	0

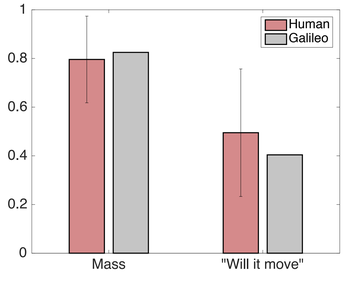


Mass estimation results

Behavior Experiments



Mean errors in outcome prediction



Accuracies on mass and stability prediction

Is this heavier?	Corr
Human vs Galileo	0.51
Human vs Truth	0.68
Galileo vs Truth	0.52
Will it move?	**Corr**
Human vs Galileo	0.56
Human vs Truth	0.42
Galileo vs Truth	0.20

Qualitative results: humans (heatmap), Galileo (orange), truth (white)

Conclusion

- Galileo: a generative model that learns physical object properties from vision
- Behavior experiments which evaluate the performance of humans and Galileo