**Motivation**

**Goal:** building a generative model to interpret physical scenes

**A Video Dataset**

101 Objects from 15 categories

**Observation**

- Humans 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.

**Scene Modeling**

**Descripive Physical Properties**
- Velocity
- Bounce
- Height
- Extended Distance

**Intrinsic Physical Object Properties**
- Coeff. Friction
- Coeff. Restitution
- Mass
- Density
- Volume

**The Galileo Model**

Initialize MCMC

Draw two physical objects

Simulated velocities $(v_{1x}, v_{1y})$

Likelihood function

Observed velocities $(v_{2x}, v_{2y})$

Tracking algorithm

Cropped images

**Experiments**

<table>
<thead>
<tr>
<th>Method</th>
<th>MSE</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle</td>
<td>0.042</td>
<td>0.77</td>
</tr>
<tr>
<td>Galileo</td>
<td>0.052</td>
<td>0.44</td>
</tr>
<tr>
<td>Uniform</td>
<td>0.081</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**Simulation results**

**Mass estimation results**

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

**Conclusion**

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